The Lebesgue Fundamental Theorem of Calculus

Generalising The Lebesgue Differentiation Theorem to Averages over Rectangles

by

M. S. Goedhart

TU Delft

The Lebesgue Fundamental Theorem of Calculus

Generalising The Lebesgue Differentiation Theorem to Averages over Rectangles

By

M.S. Goedhart

In partial fulfilment of the requirements for the degree of:

Bachelor of Science
in Applied Mathematics

at the Delft University of Technology,
to be defended publicly on [Jun 26th, 2025 at 13:00]

Supervisor: Dr. ir. E. Lorist

Thesis committee: Prof. dr. D.C. Gijswijt

Plain-Language Summary

When you take the average temperature inside a shrinking box, you expect that average to approach the temperature at the centre of the box as the box becomes smaller and smaller. Mathematicians call this idea a differentiation theorem. The integral, i.e. average, of a function returns the function's true value when the region over which we average decreases to a single point. This has been known to be true for simple shapes like balls and cubes for over a century.

But what happens if these regions are not as nicely shaped as balls or cubes? What if the regions look more like long and thin rectangles? The averaging property, as we described it earlier, then fails and the average can even blow up to infinity. In the 1930s, the mathematicians Jessen, Marcinkiewicz, and Zygmund discovered that there is a solution to this problem. Instead of requiring the function f to be integrable, we require $f(1 + \log^+ f)^{d-1}$ to be integrable. This means that we need the function to not grow too wildly; it is tamed by a logarithmic weight. With this condition, the predictable averaging behaviour also holds for rectangles. This is an important result in the mathematical field known as harmonic analysis and is even relevant in image smoothing and parts of data analysis.

This thesis revisits that result and builds it starting from an undergraduate level. Beginning with basic measure theory, we then introduce tools that can track "worst-case" local averages across regions, and we prove the necessary intermediate results. Using this, we give a complete proof of the result found by Jessen, Marcinkiewicz, and Zygmund. We also cover Zygmund's later extension of the result and his attempt to expand the theorem even further. Finally, we discuss where his conjecture breaks down and discuss how this leaves room for future research.

Abstract

This thesis provides a modern and self–contained study of integral differentiation with respect to axis–aligned rectangles. It focuses on the classical Jessen–Marcinkiewicz–Zygmund (JMZ) Theorem and the later extension by Zygmund. Throughout, our aim is to make the underlying theory and its results accessible to undergraduate–level readers. Every intermediate result is proved in full, ensuring all the essential details are appreciated.

After a brief review of the preliminaries, including measure theory and L^p, $L^{p,\infty}$ and $L(\log L)^k$ spaces, we then discuss dyadic intervals and the Hardy–Littlewood maximal operator. An important intermediate result is the Lebesgue Differentiation Theorem (LDT). We provide a proof of the LDT that avoids any covering lemmas and instead uses the properties of dyadic intervals. Both the weak–L^1 and strong L^p bounds for maximal operators are also presented. Together with the Lebesgue Differentiation Theorem, they form the basis for the main part of the thesis.

A detailed and step–by–step reconstruction of the 1935 JMZ Theorem forms the core of this thesis. The theorem extends the Lebesgue Differentiation Theorem from balls and cubes to the more general axis–aligned rectangles. We give a complete proof of the theorem and also show that the condition $f \in L(\log^+ L)^{d-1}(\mathbb{R}^d)$ is sharp.

Further, we revisit Zygmund's 1967 extension, which considers rectangles with $1 \leq k \leq d$ distinct side lengths. This naturally leads to a discussion of Zygmund's Conjecture, which tries to push the limits of Zygmund's extension of the JMZ Theorem. We analyse the counterexamples by F. Soria and G. Rey, and the examples by A. Córdoba and F. Soria. Each step in the construction of their examples and counterexamples is explained in detail. Finally, we close the thesis with a few concluding remarks and an outlook for possible future research on the conditions under which Zygmund's Conjecture holds – or fails.

Contents

Chapter 1

Introduction

The Fundamental Theorem of Calculus tells us that, in the pleasant setting of continuous functions in one variable, integration and differentiation are each other's inverses. While there is no arguing that this is an enormously useful and important result within mathematics, it is in some cases a bit limiting. In many fields of mathematics we do not deal solely with continuous functions in one variable but with measurable functions in $\mathbb{R}^d$. This is where the Lebesgue Differentiation Theorem (LDT) comes in. It states that for a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ such that $f \in L^1(\mathbb{R}^d)$,

$$\lim_{r \downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(y) - f(x)| \, \mathrm{d}y = 0 \qquad \text{for almost every} \;\; x \in \mathbb{R}^d \,, \tag{1.1}$$

where $Q_r(x)$ denotes an open cube of radius r centred at x. Instead of using cubes Q_r we could also use open balls B_r and the theorem would still hold. [1, 2, 3]

Problems start to arise when we do not average over cubes or balls, which have only a single size parameter, but average over axis–aligned rectangles instead. As it turns out, there are many functions f for which the equality in (1.1) does not hold. This discovery has led to a wide range of mathematical papers in the fields of harmonic analysis and geometric measure theory.

Jessen, Marcinkiewicz, and Zygmund published their classical paper "Note on the Differentiability of Multiple Integrals" [4] in 1935. They proved that if we impose a slightly more restrictive condition on the measurable function f, then the averages taken over arbitrary axis–aligned rectangles do converge to $f(x)$. This result is called the Jessen–Marcinkiewicz–Zygmund (JMZ) Theorem and forms the core of this thesis.

Zygmund later proved that if, instead of all axis–aligned rectangles with d distinct side lengths, we consider only rectangles with $k \leq d$ distinct side lengths, we can slightly loosen the restriction on f. [5]

When averages of integrable functions recover the original function values, differentiation theorems connect real and harmonic analysis, geometry, probability, and fields of applied mathematics. Via maximal–operators, classical results like the Lebesgue Differentiation Theorem, the Jessen–Marcinkiewicz–Zygmund Theorem, and Zygmund's extension are able to turn questions about geometric coverings and tilings into analytic convergence problems, and vice versa. This forms the basis of modern results on singular integrals, PDE's, and almost–sure limit laws. Even in more applied subjects the results are relevant. Averaging over rectangular–grids is used in many algorithms in image processing, computer vision, and data analysis. By knowing when and where averages converge or fail to converge helps improve how these grids are built and therefore advances algorithms. So while the results by Lebesgue and Jessen, Marcinkiewicz, and Zygmund are not new, they continue to be relevant today.

This thesis aims to provide a complete text on the Jessen–Marcinkiewicz–Zygmund Theorem (1935), Zygmund's Theorem (1967), and Zygmund's Conjecture (1967). Chapter 2 provides a brief review of the necessary background in measure theory and L^p–spaces. We then develop the theory of maximal functions and the Lebesgue Differentiation Theorem in Chapter 3. The core result, the Jessen–Marcinkiewicz–Zygmund theorem, is presented, followed by a proof of its sharpness in Chapter 4. We then cover Zygmund's Theorem and Zygmund's Conjecture in Chapter 5 and present some counterexamples to the conjecture. Finally, we end with some concluding remarks and an outlook for future research in Chapter 6.

Chapter 2

Essential Preliminaries

In order to be able to say and discuss anything of value about the Lebesgue Differentiation Theorem and the theorems that use its result, it is necessary to have a strong real analysis basis. In this chapter, we briefly revisit the necessary concepts. Most of the theory that is discussed should be familiar to the reader. The definitions, propositions, lemmas, and theorems that are mentioned here are found in and based on the excellent texts [1, 2, 3, 6, 7].

2.1. Measures and Integration

We start at the beginning by stating some definitions and basic theorems concerning measurable spaces, measures, measurable functions, and integration with respect to a measure.

2.1.1. σ–Algebras and Measurable Spaces

A natural place to start is with σ–algebras. First, we let $\mathcal{P}(S)$ denote the collection of all subsets of a set S, also called the power set of S.

Definition 2.1.1. *Let S be a set. A collection $\mathcal{A} \subseteq \mathcal{P}(S)$ is called a σ–algebra if*

i. $\emptyset, S \in \mathcal{A}$,

ii. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,

iii. $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$.

The pair $(S, \mathcal{A})$ is called a measurable space. The sets $A \in \mathcal{A}$ are called measurable sets.

The σ–algebras that are of interest to us in the context of integration and differentiation, and perhaps even the most important σ–algebras, are the *Borel* σ–algebras.

Definition 2.1.2 (Borel σ–algebra). *Let (M, d) be a metric space. Let $\mathcal{B}(M)$ be the σ–algebra generated by all the open sets in M, i.e. the smallest σ–algebra that contains all open sets in M, so*

$$\mathcal{B}(M) := \sigma(\{O \subseteq M : O \text{ is open}\}).$$

*The σ–algebra $\mathcal{B}(M)$ is called the **Borel σ–algebra** of M. The elements of $\mathcal{B}(M)$ are called **Borel sets**.*

2.1.2. Measures and Measure Spaces

In order to actually do something with measure theory, we need to define what a measure is. Measures provide us with a way to quantify the size of a set. This makes integration, as we define it later, possible.

Definition 2.1.3. *Let $(S, \mathcal{A})$ be a measurable space. A map $\mu : \mathcal{A} \to [0, \infty]$ is called a **measure** if*

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

ii. if for all disjoint $A_1, A_2, \ldots \in \mathcal{A}$ that satisfy $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, one has

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

*If μ is a measure, the triple $(S, \mathcal{A}, \mu)$ is called a **measure space**. If additionally $\mu(S) = 1$, then μ is called a **probability measure** and $(S, \mathcal{A}, \mu)$ is called a **probability space**.*

Remark 2.1.4. A measure space $(S, \mathcal{A}, \mu)$ is called *σ–finite* if there exists a sequence $S_n \in \mathcal{A}$ with $\mu(S_n) < \infty$ for all $n \geq 1$ such that $S = \bigcup_{n=1}^{\infty} S_n$.

Throughout this thesis we often require *completeness* of the measure space we consider to ensure that functions behave properly under integration and convergence.

Definition 2.1.5 (Completeness). *A measure space $(S, \mathcal{A}, \mu)$ is called **complete** if for every set $A \in \mathcal{A}$ with $\mu(A) = 0$ and every $E \subseteq A$, one has $E \in \mathcal{A}$ and $\mu(E) = 0$.*

The following proposition is needed when we discuss integration, specifically multiple integrals, in Subsection 2.1.4. It gives us a natural way to define a measure on σ–algebras in higher dimensions, and this new *product measure* is used in Fubini's Theorem (Thm. 2.1.15), which is introduced in Subsection 2.1.4.

Proposition 2.1.6. *Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be σ–finite measure spaces. Then there exists a unique measure $\mu \times \nu : \mathcal{A} \times \mathcal{B} \to [0, \infty]$ such that*

$$(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B), \qquad A \in \mathcal{A},\ B \in \mathcal{B},$$

and the measure space $(A \times B, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$ is σ–finite.

While general measures are certainly an interesting topic of study, for the sake of this thesis we only work with the *Lebesgue* measure λ because its translation invariance, simple behaviour under scalar multiplication, and completeness of the Lebesgue σ–algebra are crucial in the coming chapters.

Before we proceed with the Lebesgue measure and σ–algebra, we give a way to quantify the size of a particularly simple set, the half–open rectangle. For any half–open rectangle $I = (a, b] \in \mathbb{R}^d$ with $a = (\alpha_1, \ldots, \alpha_d)$ and $b = (\beta_1, \ldots, \beta_d)$ and $\alpha_j \leq \beta_j$ for $j = 1, \ldots, d$, we define its volume in the natural way by

$$|I| := \prod_{j=1}^{d} (\beta_j - \alpha_j).$$

The Lebesgue measure is constructed using this definition of the volume of rectangles in $\mathbb{R}^d$. We define the class of all Lebesgue–measurable sets in $\mathbb{R}^d$ as the Lebesgue σ–algebra denoted by $\mathcal{L}(\mathbb{R}^d)$. The following theorem summarises the most significant properties of the Lebesgue measure. Here uniqueness is due to Carathéodory's Extension Theorem. The reader is referred to [3, Chap. 1] for a proof of the completeness of $\mathcal{L}(\mathbb{R}^d)$ and details on the uniqueness of the measure.

Theorem 2.1.7 (Lebesgue Measure on $\mathcal{L}(\mathbb{R}^d)$). *There exists a unique measure λ on $(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d))$ such that $\lambda(I) = |I|$ for all half–open rectangles I. Moreover, for all $h \in \mathbb{R}^d$ and $A \in \mathcal{L}(\mathbb{R}^d)$, $\lambda(A + h) = \lambda(A)$ and $\lambda(hA) = |h|\,\lambda(A)$. Here, $A + h := \{x + h : x \in A\}$ and $hA := \{hx : x \in A\}$.*

2.1.3. Measurable Functions

Before we can move on to integration with respect to a measure, we need to cover *measurable functions*. Measurable functions are those functions $f : S \to \mathbb{R}$ for which the range can be discretised by a measure. Without this property, integration of f would not make sense.

Definition 2.1.8. *Let $(S, \mathcal{A})$ and $(T, \mathcal{B})$ be two measurable spaces. A function $f : S \to T$ is called* ***measurable*** *if for each $B \in \mathcal{B}$, we have $f^{-1}(B) \in \mathcal{A}$. In words, pre–images of measurable sets are measurable.*

In many estimates, we allow functions to take the values ∞ and $-\infty$. We therefore extend $\mathbb{R}$ to include these values and denote

$$\bar{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}.$$

We can naturally extend the Lebesgue σ–algebra, $\mathcal{L}(\mathbb{R})$, to $\bar{\mathbb{R}}$ for functions $f : S \to \bar{\mathbb{R}}$. The spaces $\mathbb{R}$, $\bar{\mathbb{R}}$, and $[0, \infty]$ are always equipped with the corresponding Lebesgue σ–algebras $\mathcal{L}(\mathbb{R})$, $\mathcal{L}(\bar{\mathbb{R}})$, and $\mathcal{L}([0, \infty])$ unless specified otherwise.

Remark 2.1.9. Algebraic combinations, suprema, infima, limits superior, and limits inferior all preserve measurability. The reader is referred to [6, 7], or any other undergraduate textbook on real analysis, for the proofs.

Measurable functions need not be real–valued. We can have complex–valued functions as well. Any function $f : S \to \mathbb{C}$ can be written as $f = u + iv$, where $u, v : S \to \mathbb{R}$ are the real and imaginary parts. It immediately follows that a function $f : S \to \mathbb{C}$ is measurable *if and only if* the functions u and v are measurable.

Just like we extended $\mathbb{R}$, we can extend $\mathbb{C}$ to allow functions to take the value ∞. This is denoted as

$$\bar{\mathbb{C}} := \mathbb{C} \cup \{\infty\}.$$

All properties of measurable functions carry over naturally when using the decomposition into real and imaginary parts.

An important theorem regarding measurable functions is *Lusin's Theorem*. It forms one of the main pillars of measure–theoretic real analysis and creates a bridge between measurable and continuous functions. The intuition that "every measurable function is nearly continuous" is confirmed by this theorem.

Theorem 2.1.10 (Lusin's Theorem). *Let $(S, \mathcal{A}, \lambda)$ be a measure space where λ denotes the Lebesgue measure and $S \subseteq \mathbb{R}^d$. If $f : S \to \mathbb{C}$ is a measurable function and S has finite measure, then for any $\varepsilon > 0$, there exists a continuous function φ such that $f = \varphi$ except on a set of measure strictly less than ε.*

We come back to Lusin's Theorem after we cover L^p–spaces in Section 2.2 to prove a very useful approximation result that we use when proving the Jessen–Marcinkiewicz–Zygmund Theorem.

2.1.4. Integration

Integration theory is built from approximating measurable functions by simple functions. These simple functions are a linear combination of the most basic measurable functions: the characteristic functions. Let $(S, \mathcal{A})$ be a measurable space, then for $A \subseteq S$, the characteristic function is defined as

$$\mathbb{1}_A(s) := \begin{cases} 1 & s \in A \\ 0 & s \notin A. \end{cases}$$

Note that $\mathbb{1}_A$ is measurable if and only if $A \in \mathcal{A}$.

We do not go into the full details of constructing integration theory in this way. The important things to mention are the definition of the Lebesgue integral for simple functions, how this extends to measurable functions, and some basic properties.

Definition 2.1.11. *Let $(S, \mathcal{A})$ be a measurable space and let $f : S \to [0, \infty)$ be a simple function. Let $x_1, \ldots, x_n \in [0, \infty)$ and let $(A_k)_{k=1}^n \in \mathcal{A}$ be disjoint sets such that*

$$f = \sum_{k=1}^{n} x_k \cdot \mathbb{1}_{A_k}. \tag{2.1}$$

For $E \in \mathcal{A}$ we define

$$\int_E f \, \mathrm{d}\mu := \sum_{k=1}^{n} x_k \cdot \mu(E \cap A_k),$$

*which is called the **Lebesgue integral** of f over the set E.*

The Lebesgue integral of a general measurable function is defined by using the decomposition into positive and negative parts of the function. The measurable function $f : S \to \bar{\mathbb{R}}$ is split into two *non–negative* measurable functions as $f = f^+ - f^-$, where $f^+ := \max\{f, 0\}$ and $f^- := \max\{-f, 0\}$.

Definition 2.1.12 (Lebesgue Integral). *The Lebesgue integral of the non–negative measurable functions $f^\pm : S \to [0, \infty]$ over $E \in \mathcal{A}$ is defined as*

$$\int_E f^\pm \, \mathrm{d}\mu := \lim_{n\to\infty} \int_E f_n^\pm \, \mathrm{d}\mu \quad \text{in} \quad [0, \infty]$$

where $(f_n^+)_{n\geq1}$ and $(f_n^-)_{n\geq1}$ are sequences of simple functions with $0 \leq f_n^+ \uparrow f^+$ and $0 \leq f_n^- \uparrow f^-$.

*A measurable function $f : S \to \bar{\mathbb{R}}$ is called **integrable** when both $\int_S f^+ \, \mathrm{d}\mu$ and $\int_S f^- \, \mathrm{d}\mu$ are finite. In this case, the **Lebesgue integral** of f over $E \in \mathcal{A}$ is defined as*

$$\int_E f \, \mathrm{d}\mu := \int_E f^+ \, \mathrm{d}\mu - \int_E f^- \, \mathrm{d}\mu.$$

From the two given definitions, the following basic properties are readily deduced. For a more detailed discussion on the construction of the Lebesgue integral and proofs of the properties mentioned here, we refer the reader to [6, 7] or similar textbooks.

Lemma 2.1.13. *Let $(S, \mathcal{A})$ be a measurable space and let $f, g : S \to \bar{\mathbb{R}}$ be integrable functions. The following hold:*

i. For all $E \in \mathcal{A}$, $\int_E f \, \mathrm{d}\mu = \int_S \mathbb{1}_{\mathbb{E}} \cdot f \, \mathrm{d}\mu$,

ii. If $E \in \mathcal{A}$ and $f \leq g$ on E, then $\int_E f \, \mathrm{d}\mu \leq \int_E g \, \mathrm{d}\mu$,

iii. For all $E \in \mathcal{A}$ and $\alpha, \beta \in [0, \infty)$, $\int_E \alpha f + \beta g \, \mathrm{d}\mu = \alpha \int_E f \, \mathrm{d}\mu + \beta \int_E g \, \mathrm{d}\mu$,

iv. For all disjoint sets $E, F \in \mathcal{A}$, $\int_{E\cup F} f \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu + \int_F f \, \mathrm{d}\mu$,

v. f is integrable if and only if $|f|$ is integrable, in this case for each $E \in \mathcal{A}$, $|\int_E f \, \mathrm{d}\mu| \leq \int_E |f| \, \mathrm{d}\mu$.

Remark 2.1.14. For any continuous function f, the Lebesgue integral with respect to the Lebesgue measure λ coincides with the Riemann integral. Moreover, any Riemann–integrable function is also integrable with respect to the Lebesgue measure. This also holds for *improper* Riemann integrals in the case that $\int_{-\infty}^{\infty} f(x) \, \mathrm{d}x < \infty$.

We end this section by giving a very useful result: *Fubini's Theorem.* This theorem provides us with the conditions under which we can compute a double integral using iterated integrals. It also states that, under those conditions, the order in which we use iterated integrals is irrelevant.

Theorem 2.1.15 (Fubini's Theorem). *Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be σ–finite measure spaces. If $f : S \times T \to \mathbb{R}$ is integrable, then*

$$\int_{S\times T} f \, \mathrm{d}(\mu \times \nu) = \int_T \int_S f \, \mathrm{d}\mu \, \mathrm{d}\nu = \int_S \int_T f \, \mathrm{d}\nu \, \mathrm{d}\mu.$$

2.2. L^p–, $L^{p,\infty}$–, and $L(\log^+ L)^k$–spaces

When working with Lebesgue integrals and differentiation, we want to constrain ourselves to spaces of functions with relatively nice properties. This is where the L^p–, $L^{p,\infty}$–, and $L(\log^+ L)^k$–spaces come in. Throughout this section, $(S, \mathcal{A}, \mu)$ denotes a fixed σ–finite, complete measure space in the sense of Definition 2.1.8 and $p \in (0, \infty]$.

2.2.1. L^p–spaces

Definition 2.2.1. *For $f : S \to \mathbb{C}$ a measurable function and $p \in (0, \infty]$, we define*

$$||f||_p := ||f||_{L^p(S)} = \left(\int_S |f|^p \, \mathrm{d}\mu\right)^{1/p}, \qquad p \in (0, \infty),$$
$$||f||_\infty := ||f||_{L^\infty(S)} = \operatorname*{ess\,sup}_{s\in S} |f(s)|, \qquad p = \infty,$$

where $\operatorname{ess\,sup}_{s\in S}|f(s)| = \inf\{r \geq 0 : |f| \leq r \text{ a.e.}\}$.

We define $L^p(S)$ as the space of all complex–valued measurable functions $f : S \to \mathbb{C}$ for which $||f||_p < \infty$. Within $L^p(S)$–spaces, the functions that are *almost everywhere* equal are identified. A function f is called *integrable* if $f \in L^1(S)$.

Given some $p \in [1, \infty]$, we define $p' \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{p'} = 1$. This p' is called the *Hölder conjugate* of p. The Hölder conjugate gives rise to an important inequality.

Proposition 2.2.2 (Hölder's Inequality). *Let $p \in [1, \infty]$, then for all measurable $f, g : S \to \mathbb{C}$*

$$||fg||_1 \leq ||f||_p ||g||_{p'}.$$

Moreover, if $(q_j)_{j=1}^n$, $q \in (0, \infty]$ satisfy $\sum_{j=1}^n \frac{1}{q_j} = \frac{1}{q}$ and $(f_j)_{j\geq 1}$ are measurable functions, then

$$||f_1 \cdots f_n||_q \leq ||f_1||_{q_1} \cdots ||f_n||_{q_n}.$$

In general, not much can be said about how two L^p–spaces are related: we have $L^p \not\subseteq L^q$ and $L^q \not\subseteq L^p$. In the case that $\mu(S) < \infty$ or S is a discrete set with the counting measure, we can say how the spaces are related. A useful result that follows from Hölder's Inequality is formulated in the following proposition.

Proposition 2.2.3. *Let $0 < p < q \leq \infty$. If $\mu(S) < \infty$, then $L^q(S) \subseteq L^p(S)$ and for all measurable $f : S \to \mathbb{C}$*

$$||f||_p \leq \mu(S)^{\frac{1}{p} - \frac{1}{q}} ||f||_q.$$

Proof. We use Hölder's Inequality with functions f and g where $g(s) := 1$. The case where $q = \infty$ is straightforward. By assumption, we have $||f||_\infty < \infty$. Note that $|f(s)| \leq ||f||_\infty$ for almost every $s \in S$. Integrating gives

$$||f||_p^p = \int_S |f(s)|^p \,\mathrm{d}\mu(s) \leq ||f||_\infty^p \mu(S).$$

If $q < \infty$, then

$$||f||_p^p = \int_S |f(s)|^p \cdot 1 \,\mathrm{d}\mu(s) \leq \left(\int_S |f(s)|^q \,\mathrm{d}\mu(s) \right)^{p/q} \left(\int_S 1 \,\mathrm{d}\mu(s) \right)^{(q-p)/q} = ||f||_q^p \mu(S)^{(q-p)/q}. \qquad \square$$

Another famous result we use is Jensen's Inequality. This inequality allows us to relate the convex function of an integral to the integral of the convex function. Hölder's Inequality is a special case of Jensen's Inequality, where φ is taken to be $\varphi(t) = t^p$.

Theorem 2.2.4 (Jensen's Inequality). *Let $(S, \mathcal{A}, \mu)$ be a σ–finite measure space. Let $f : S \to \mathbf{R}$ be a measurable function and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function. Then*

$$\varphi\left(\frac{1}{\mu(S)} \int_S f \,\mathrm{d}\mu \right) \leq \frac{1}{\mu(S)} \int_S \varphi(f) \,\mathrm{d}\mu.$$

The next inequality deals with functions in two variables. The *Minkowski Inequality* is the triangle inequality generalised to L^p–spaces. The proof is based on Fubini's Theorem (Thm. 2.1.15) and Hölder's Inequality (Prop. 2.2.2).

Proposition 2.2.5 (Minkowski's Integral Inequality). *Let $p \in [1, \infty]$ and let $(S, \mathcal{A}, \mu)$, $(T, \mathcal{B}, \nu)$ be σ–finite measure spaces. Suppose that $f : S \times T \to \mathbb{C}$ is measurable. Then*

$$\left(\int_S \left[\int_T |f(s,t)| \,\mathrm{d}\nu(t) \right]^p \mathrm{d}\mu(s) \right)^{1/p} \leq \int_T \left(\int_S |f(s,t)|^p \,\mathrm{d}\mu(s) \right)^{1/p} \mathrm{d}\nu(t).$$

Now that we have covered the essentials of L^p–spaces, we briefly circle back to Lusin's Theorem (Thm. 2.1.10). We can use this theorem to prove an approximation result that we need in Chapter 4.

Lemma 2.2.6. *Let $f \in L^1(S)$ and $\varepsilon > 0$, then there exists a continuous function $\varphi \in L^1(S)$ such that*

$$\int_S |f(s) - \varphi(s)| \,\mathrm{d}\mu(s) < \varepsilon\,.$$

Proof. We prove this only for real–valued functions f. In case that f is complex–valued, we apply the result to the real and imaginary parts separately.

Note that $f \cdot \mathbb{1}_{\{|f| \geq n\}} \to 0$ as $n \to \infty$ and that $|f \cdot \mathbb{1}_{\{|f| \geq n\}}| \leq |f|$. It follows from the Dominated Convergence Theorem that

$$\lim_{n \to \infty} \int_{f \cdot \mathbb{1}_{\{|f| \geq n\}}} |f(s)| \,\mathrm{d}\mu(s) = 0.$$

Let $\delta > 0$ be arbitrary and choose some N such that $\int_{f \cdot \mathbb{1}_{\{|f| \geq n\}}} |f| \,\mathrm{d}\mu < \delta$ for all $n \geq N$. Now define

$$f_n(x) := \begin{cases} f(x)\,, & |f(x)| \leq n\,, \\ n\,, & f(x) > n\,, \\ -n\,, & f(x) < -n\,. \end{cases}$$

Let $\varphi : S \to \mathbb{R}$ be the continuous function such that $f = \varphi$ on $E \subseteq S$ and $\lambda(S \setminus E) < \delta/n$. We know such a function exists by Lusin's Theorem.

We may assume that $|\varphi| \leq n$ almost everywhere, otherwise we redefine it by setting $\varphi(x) = n$ where earlier $\varphi(x) > n$ and setting $\varphi(x) = -n$ where earlier $\varphi(x) < -n$. Now,

$$\begin{aligned}\int_S |f(s) - \varphi(s)| \, \mathrm{d}\mu(s) &\leq \int_S |f - f_n| \, \mathrm{d}\mu + \int_S |f_n - \varphi| \, \mathrm{d}\mu \\ &= \int_{\{|f| \geq n\}} |f - f_n| \, \mathrm{d}\mu + \int_{S \setminus E} |f_n - \varphi| \, \mathrm{d}\mu \\ &\leq \int_{\{|f| \geq n\}} |f| \, \mathrm{d}\mu + 2n \cdot \mu(S \setminus E) \\ &< \delta + 2\delta = 3\delta.\end{aligned}$$

Setting $\varepsilon := \frac{1}{3}\delta$ proves the lemma. □

2.2.2. $L^{p,\infty}$–spaces

From L^p-spaces we move on to $L^{p,\infty}$–spaces, the *weak* L^p–spaces. These spaces have a less restrictive definition and, therefore, contain more functions.

Definition 2.2.7. *For $f : S \to \mathbb{C}$ a measurable function and $p \in (0, \infty)$, we define*

$$||f||_{p,\infty} = ||f||_{L^{p,\infty}(S)} = \sup_{t>0} \; t\mu(\{|f| > t\})^{1/p}.$$

Remark 2.2.8. $||f||_{p,\infty}$ is the smallest constant C such that $\mu(\{|f| > t\}) \leq \frac{C^p}{t^p}$ for all $t > 0$.

It should be noted that while convention dictates that we write $||\cdot||_{p,\infty}$, this does not actually define a norm; the triangle inequality fails.

The $L^{p,\infty}$–space is defined as the space of all complex–valued measurable functions $f : S \to \mathbb{C}$ for which $||f||_{p,\infty} < \infty$. We again identify functions that are equal almost everywhere.

An important result in both real– and harmonic analysis is *Chebyshev's Inequality.* It relates the L^p and $L^{p,\infty}$ (quasi–)norms.

Proposition 2.2.9 (Chebyshev's Inequality). *Let $p \in (0, \infty)$. For every measurable function $f : S \to \mathbb{C}$, we have*

$$||f||_{p,\infty} \leq ||f||_p.$$

Proof. Let $t > 0$ be arbitrary, then

$$t^p \mu(\{|f| > t\}) = \int_S t^p \mathbb{1}_{\{|f|>t\}} \, \mathrm{d}\mu \leq \int_S |f(s)|^p \mathbb{1}_{\{|f|>t\}} \, \mathrm{d}\mu(s) \leq ||f||_p^p.$$

In particular, $\sup_{t>0} t^p \mu(\{|f| > t\}) \leq ||f||_p^p$ and thus $||f||_{p,\infty} \leq ||f||_p$. □

2.2.3. $L(\log^+ L)^k$–spaces

To bridge the gap between $L^1(S)$– and $L^p(S)$–spaces, the $L(\log^+ L)^k(S)$–spaces are introduced. These spaces are used often when discussing the Jessen–Marcinkiewicz–Zygmund Theorem in Chapter 4. We would expect that $L(\log^+ L)^k(S)$ is defined as the space of all complex–valued measurable functions $f : S \to \mathbb{C}$ for which

$$\int_S |f| (\log^+ |f|)^k \, \mathrm{d}\mu < \infty \, ,$$

but this is not the case. If we were to use that definition, then $L(\log^+ L)^k$ would not be a vector space; $\int_S |f|(\log^+|f|)^k \, \mathrm{d}\mu$ is not a norm. It is easily checked that not even multiplication with a scalar works as we would expect. To work around this, it is common to use the following definition.

Definition 2.2.10. *For $f : S \to \mathbb{C}$ a measurable function and $k > 0$, we define $L(\log^+ L)^k(S)$ as the space of all f for which*

$$\int_S |f|(1 + \log^+|f|)^k \, \mathrm{d}\mu < \infty \,,$$

where $\log^+ f := \max(\log^+, 0)$.

Just like in L^p– and $L^{p,\infty}$–spaces, we identify functions that are equal almost everywhere.

We started this subsection by stating that $L(\log^+ L)^k$ spaces bridge the gap between the L^1 and L^p spaces. What we mean by this becomes clear when we consider a measure space with finite measure. We already saw in Proposition 2.2.3 that if $\mu(S) < \infty$, then $L^q(S) \subseteq L^p(S)$ when $0 < p \leq q \leq \infty$. This leads us to the following lemma.

Lemma 2.2.11. *Let $0 < k_1 < k_2 < \infty$ and $1 < p \leq \infty$. If $\mu(S) < \infty$, then $L^p(S) \subseteq L(\log^+ L)^{k_2}(S) \subseteq L(\log^+ L)^{k_1}(S) \subseteq L^1(S)$.*

Proof. We start by proving the inclusion $L^p(S) \subseteq L(\log^+ L)^{k_2}(S)$. Let $f \in L^p(S)$ with $1 < p \leq \infty$. We split the integral over the sets $\{|f| \leq 1\}$ and $\{|f| > 1\}$. On the set $\{|f| \leq 1\}$, we have $\log^+|f(x)| = 0$, so $|f(x)|(1 + \log^+|f(x)|)^{k_2} = |f(x)|$. On the set $\{|f| > 1\}$, we have $\log^+|f(x)| = \log|f(x)| \leq |f(x)|$, so $(1 + \log^+|f(x)|)^{k_2} \leq C\,|f(x)|^{\varepsilon}$ for any $\varepsilon > 0$ and some constant C depending on k_2 and ε. Thus,

$$\begin{aligned}
\int_S |f(y)|(1 + \log^+|f(y)|)^{k_2} \, \mathrm{d}y &= \int_{\{|f| \leq 1\}} |f(y)| \, \mathrm{d}y + \int_{\{|f| > 1\}} |f(y)|(1 + \log^+|f(y)|)^{k_2} \, \mathrm{d}y \\
&\leq \int_{\{|f| \leq 1\}} |f(y)| \, \mathrm{d}y + \int_{\{|f| > 1\}} C|f(y)|^{1+\varepsilon} \, \mathrm{d}y \\
&\leq \int_{\{|f| \leq 1\}} |f(y)|^p \, \mathrm{d}y + \int_{\{|f| > 1\}} C|f(y)|^p \, \mathrm{d}y < \infty.
\end{aligned}$$

In the last inequality, we used that we can choose $\varepsilon > 0$ such that $1 + \varepsilon = p$.

The inclusion $L(\log^+ L)^{k_2}(S) \subseteq L(\log^+ L)^{k_1}(S)$ follows immediately from the fact that $(1 + \log^+|f(x)|)^n \leq (1 + \log^+|f(x)|)^m$ for all $0 < n < m$. Let $f \in L(\log^+ L)^{k_2}(S)$, then

$$\int_S |f(y)|(1 + \log^+|f(y)|)^{k_1} \, \mathrm{d}y \leq \int_S |f|(1 + \log^+|f|)^{k_2} \, \mathrm{d}y < \infty$$

and thus $f \in L(\log^+ L)^{k_1}(S)$.

Lastly, let $f \in L(\log^+ L)^{k_1}(S)$. Note that $(1 + \log^+|f(x)|)^{k_1} \geq 1$ for all $f(x)$. It immediately follows that

$$\int_S |f(y)| \, \mathrm{d}y \leq \int_S |f(y)|(1 + \log^+|f(y)|)^{k_1} \, \mathrm{d}y < \infty,$$

and thus $f \in L^1(S)$.

Together, this shows that $L^p(S) \subseteq L(\log^+ L)^{k_2}(S) \subseteq L(\log^+ L)^{k_1}(S) \subseteq L^1(S)$. □

2.3. The Fundamental Theorem of Calculus

In the last section of this chapter, we very briefly recall the two Fundamental Theorems of Calculus. We expect that the reader is very familiar with these theorems, but we state them for completeness. Chapter 3 builds on these notions to form the theory needed for the Jessen–Marcinkiewicz–Zygmund Theorem.

Theorem 2.3.1 (Fundamental Theorem of Calculus I). *Let f be an integrable function on $[a, b]$. For each $x \in [a, b]$, define*

$$F(x) := \int_a^x f(y)\,\mathrm{d}y.$$

Then F is uniformly continuous on $[a, b]$. Moreover, if f is continuous at $c \in [a, b]$, then F is differentiable at c and

$$F'(c) = f(c).$$

Theorem 2.3.2 (Fundamental Theorem of Calculus II). *Let f be differentiable on $[a, b]$ and f' integrable on $[a, b]$, then*

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a).$$

We can rewrite this as follows. For a continuous function $f : \mathbb{R} \to \mathbb{C}$ and $x_0 \in \mathbb{R}$, define

$$F(x) := \int_{x_0}^x f(y)\,\mathrm{d}y, \quad x \in \mathbb{R}.$$

Then for $x \in \mathbb{R}$, by the Fundamental Theorem of Calculus, we have

$$F'(x) = \lim_{r \downarrow 0} \frac{F(x + r) - F(x - r)}{2r} = \lim_{r \downarrow 0} \frac{1}{2r} \int_{x-r}^{x+r} f(y)\,\mathrm{d}y = f(x).$$

This statement generalises to functions in $\mathbb{R}^d$. If $f : \mathbb{R}^d \to \mathbb{C}$ is continuous, then for every $x \in \mathbb{R}^d$

$$\lim_{r \downarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|\,\mathrm{d}y = 0.$$

Here $B_r(x)$ denotes the open ball centred at x with radius r and $|B_r(x)|$ denotes its Lebesgue measure.

From this formulation, the question arises if we can say something similar for more general functions than just continuous ones. The answer to this is positive and is formulated in the Lebesgue Differentiation Theorem, which is discussed in Chapter 3.

Chapter 3

Maximal Bounds and Integral Differentiation

Before proving the Jessen–Marcinkiewicz–Zygmund Theorem, we need some additional theory to the preliminaries discussed in Chapter 2. The main result in this chapter is the Lebesgue Differentiation Theorem. To prove this, we introduce the Hardy–Littlewood maximal operator and a useful bound [1, 2]. This bound relies on geometric arguments, which we build on the structure of dyadic cubes. Dyadic cubes have a few very useful properties that we exploit, particularly their nesting and disjointness [8]. Throughout this chapter, the Lebesgue measure of a set $A \subseteq \mathbb{R}^d$ is denoted by $|A|$.

3.1. Dyadic Cubes

We start this chapter by introducing the dyadic cubes, a very useful tool in different areas of analysis. Dyadic cubes slice $\mathbb{R}^d$ into axis–aligned cubes whose side lengths are integer powers of two. By doing this, a discrete and nested grid is obtained that removes the need for various covering lemmas that would normally be present in some proofs.

Definition 3.1.1. *A dyadic cube in $\mathbb{R}^d$ is a set of the form*

$$R^{(k,m)} = \left[\frac{m_1}{2^k}, \frac{m_1+1}{2^k}\right) \times \dots \times \left[\frac{m_d}{2^k}, \frac{m_d+1}{2^k}\right),$$

where $k, m_1, \dots, m_d \in \mathbb{Z}$. We denote $\mathcal{D}$ as the collection of all dyadic cubes and call this a dyadic grid. We write $\mathcal{D} = \bigcup_{k\in\mathbb{Z}} \mathcal{D}_k$ where $\mathcal{D}_k := \{R^{(k,m)} : m \in \mathbb{Z}\}$.

From the definition, it is easily seen that for any two dyadic cubes, one is contained in the other, or they are disjoint. Also, for any $k \in \mathbb{Z}$, the subcollection $\mathcal{D}_k$ partitions $\mathbb{R}^d$. This leads to the first result concerning dyadic cubes. [8]

Lemma 3.1.2. *Let $R_1, \dots, R_n$ be a finite collection of dyadic cubes in $\mathbb{R}^d$. Then there exists a subcollection of **disjoint** cubes $R_{k_1}, \dots, R_{k_m}$ such that*

$$R_1 \cup \dots \cup R_n = R_{k_1} \cup \dots \cup R_{k_m}.$$

Proof. Let R_{k_i} denote those cubes in $R_1, \dots, R_n$ that are not contained by any other cubes in the collection, i.e. $R_{k_i} \not\subset R_j$ for all $j = 1, \dots, n$. The nesting property guarantees that these R_{k_i} are disjoint and their union covers the union of $R_1, \dots, R_n$. □

These nested and partitioning properties make dyadic cubes particularly useful. Any arbitrary subset of Euclidean space can be replaced by some union of dyadic cubes that cover this subset. In order to move from a Euclidean subset to a dyadic grid that covers it, we make use of the so–called One–Third Trick. We first state it in the one–dimensional setting and then generalise it to $\mathbb{R}^d$ for any $d \geq 1$. [9]

Lemma 3.1.3. *Let $\mathcal{D}^{(0)}, \mathcal{D}^{(\frac{1}{3})}, \mathcal{D}^{(\frac{2}{3})}$ be three shifted dyadic grids on $\mathbb{R}$ defined by*

$$\mathcal{D}^{(\sigma)} := \left\{\left[\frac{m+(-1)^k\sigma}{2^k}, \frac{m+(-1)^k\sigma+1}{2^k}\right) : k, m \in \mathbb{Z}\right\},\ \sigma \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}.$$

For every bounded, half–open interval $I \subset \mathbb{R}$, there exists a shift $\sigma \in \{0, \frac{1}{3}, \frac{2}{3}\}$ and a dyadic interval $J \in \mathcal{D}^{(\sigma)}$ such that

$$I \subset J \text{ and } |I| \leq |J| \leq 3|I|.$$

Proof. Take $I = [a, b)$ with length $\ell = b - a$. Let $k \in \mathbb{Z}$ be the unique integer such that $2^{-(k+1)} \leq \ell < 2^{-k}$. Consider the three intervals of length 2^{-k} in the three grids with left endpoints

$$2^{-k}\left(\lfloor 2^k a - \sigma \rfloor + \sigma\right) \text{ where } \sigma \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}.$$

Exactly one of those three intervals contains the left endpoint a of I, we denote this interval by J. Because $|J| = 2^{-k} \leq 2\ell$ and the right endpoint of I lies a distance strictly less than ℓ from a, we must have $I \subset J$ and $|J| \leq 3|I|$. □

The one–dimensional version of the One–Third Trick easily generalises to dyadic cubes in $\mathbb{R}^d$ by writing the axis–aligned cube R as a product of bounded intervals I_i and applying Lemma 3.1.3 to each of these intervals.

Lemma 3.1.4 (One–Third Trick). *Let $\mathcal{D}^{(\sigma)}$ denote the shifted dyadic grids on $\mathbb{R}^d$ defined by*

$$\mathcal{D}^{(\boldsymbol{\sigma})} := \left\{\left[\frac{m+\sigma_1}{2^k}, \frac{m+\sigma_1+1}{2^k}\right) \times \ldots \times \left[\frac{m+\sigma_d}{2^k}, \frac{m+\sigma_d+1}{2^k}\right) : k, m \in \mathbb{Z}\right\}, \ \boldsymbol{\sigma} \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}^d$$

For every axis–aligned cube $Q \subseteq \mathbb{R}^d$, there exists a shift vector $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_d) \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$ and a dyadic cube $R \in \mathcal{D}^{(\boldsymbol{\sigma})}$, such that

$$Q \subset R \qquad \text{and} \qquad |Q| \leq |R| \leq 3^d |Q|.$$

3.2. Hardy–Littlewood Maximal Operator

With the dyadic structure, we have a strong geometric basis for the differentiation results that follow in the next section and chapters. To connect geometry and differentiation, an analytic concept, we introduce the Hardy–Littlewood maximal operator. This operator and its weak–L^1 bound form the foundation of the Lebesgue Differentiation Theorem and therefore of the Jessen–Marcinkiewicz–Zygmund Theorem. We define the Hardy–Littlewood operator in the space of *locally L^p–integrable functions.* For open $\Omega \subseteq \mathbb{R}^d$ and $p \in (0, \infty)$, we define this space as

$$L^p_{\text{loc}}(\Omega) := \{f : \Omega \to \mathbb{C} \ : f \text{ is measurable and for all compact } K \subseteq \Omega, \ \int_K |f| \, \mathrm{d}x < \infty\}.$$

Definition 3.2.1. *For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ define $Mf : \mathbb{R}^d \to [0, \infty]$ by*

$$Mf(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| \ \mathrm{d}y,$$

where the supremum is taken over all open, axis–aligned cubes $Q \subseteq \mathbb{R}^d$ that contain x.

Remark 3.2.2. The Hardy–Littlewood maximal operator can just as well be defined by considering open balls B instead of open, axis–aligned cubes Q. The definitions are equivalent: all inequalities and other properties carry over.

Remark 3.2.3. It is easily shown that $Mf > 0$ a.e. if $||f||_1 > 0$ and that Mf is measurable.

In order to use the Hardy–Littlewood operator, we need to know how it behaves. The *weak* L^1–boundedness of the Hardy–Littlewood operator gives an upper bound that is essential in the proofs of many results leading up to the JMZ Theorem. The proof that we give is based on dyadic cubes, but the boundedness could also be proved by using a covering lemma like Vitali's. Such an alternative proof can be found in [1, p. 24-26] or [2, p. 14-17].

Theorem 3.2.4 (Weak L^1–boundedness). *For all $f \in L^1(\mathbb{R}^d)$ and $t > 0$ one has*

$$|\{Mf > t\}| \leq \frac{3^d}{t} \int_{\{Mf>t\}} |f(y)| \, \mathrm{d}y,$$

and thus $||Mf||_{1,\infty} \leq ||f||_1$.

Proof. We first prove the bound for the *dyadic* maximal operator. The proof is very straightforward and the dyadic bound can then be used to prove the bound for the standard maximal operator.

Take the standard dyadic grid $\mathcal{D} := \prod_{i=1}^d \left[\frac{k_i}{2^n}, \frac{k_i+1}{2^n}\right)$ where $n, k_i \in \mathbb{Z}$. The dyadic maximal operator is then

$$M_{\mathcal{D}} f(x) := \sup_{\substack{R \in \mathcal{D} \\ R \ni x}} \frac{1}{|R|} \int_R |f(y)| \, \mathrm{d}y$$

where R is a cube in the grid $\mathcal{D}$. Fix $t > 0$ and define $E_t := \{M_{\mathcal{D}} f(x) > t\}$.

Let $\mathcal{C}_t := \{R \in \mathcal{D} : \frac{1}{|R|} \int_R f(y) \, \mathrm{d}y > t\}$ be a collection of dyadic cubes. By Lemma 3.1.2, there exists a subcollection $\mathcal{F}_t \subseteq \mathcal{C}_t$ of disjoint dyadic cubes such that $E_t = \bigcup_{R \in \mathcal{C}_t} R = \bigcup_{R \in \mathcal{F}_t} R$. We estimate the measure of E_t by

$$|E_t| = \sum_{R \in \mathcal{F}_t} |R| \leq \frac{1}{t} \sum_{R \in \mathcal{F}_t} \int_R f(y) \, \mathrm{d}y \leq \frac{1}{t} \int_{E_t} f(y) \, \mathrm{d}y.$$

Thus,

$$|\{M_{\mathcal{D}} f > t\}| \leq \frac{1}{t} \int_{\{M_{\mathcal{D}} f > t\}} |f(y)| \, \mathrm{d}y.$$

The weak L^1–boundedness is easily proved for the *dyadic* Hardy–Littlewood maximal operator because of the properties of dyadic cubes. For the standard maximal operator we need some extra steps to find a bound. This is where the One–Third Trick (Lem. 3.1.4) comes in. We can write

$$Mf(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| \, \mathrm{d}y \leq 3^d \max_{\boldsymbol{\sigma} \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} M_{\mathcal{D}^{(\boldsymbol{\sigma})}} f(x),$$

where $\mathcal{D}^{(\boldsymbol{\sigma})}$ is the shifted dyadic grid. Since we already have the bound $|\{M_{\mathcal{D}} f(x) > t\}| \leq \frac{1}{t} ||f||_{L^1}$ for the dyadic maximal operator, it follows that

$$|\{Mf(x) > t\}| \leq 3^d \left| \left\{ \max_{\boldsymbol{\sigma} \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} M_{\mathcal{D}^{(\boldsymbol{\sigma})}} f(x) > t \right\} \right| \leq \frac{3^d}{t} ||f||_1. \qquad \square$$

The weak L^1–boundedness of Mf is used directly in the proof of the Lebesgue Differentiation Theorem. For the proof of the Jessen–Marcinkiewicz–Zygmund Theorem we need an additional result: the L^p-boundedness of the Hardy–Littlewood operator. In the proof of the L^p–bound we use the identity $||f||_p^p = p \int_0^\infty p t^{p-1} \mu(\{|f| > t\}) \, \mathrm{d}t$, which follows directly from Fubini's Theorem (Thm. 2.1.15) and holds for all measurable functions provided that the measure space is σ–finite and $p \in (0, \infty)$.

Theorem 3.2.5 (L^p–boundedness). *Let $p \in (1, \infty]$. Then for all $f \in L^p(\mathbb{R}^d)$,*

$$||Mf||_p \leq 3^d p' ||f||_p,$$

where p' is the Hölder conjugate of p.

Proof. The proof for $p = \infty$ is almost trivial: for all $x \in \mathbb{R}^d$, $Mf(x) \leq \sup_{Q \ni x} \frac{1}{|Q|} \int_Q ||f||_\infty \, \mathrm{d}y = ||f||_\infty$ and thus $||Mf||_\infty \leq ||f||_\infty$. Now let $p \in (1, \infty)$ and suppose that $f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \subseteq L^p(\mathbb{R}^d)$. By Theorem 3.2.4, $Mf \in L^{1,\infty}(\mathbb{R}^d)$ and by the first part of the proof, $Mf \in L^\infty(\mathbb{R}^d)$. By definition of these spaces, it follows that there exists an $r_0 \in \mathbb{R}$ such that $|Mf| \leq r_0$ a.e. so $|\{|Mf| > t\}| = 0$ for all $t \geq r_0$. We then write,

$$\int_0^\infty pt^{p-1} |\{|Mf| > t\}| \, \mathrm{d}t = \int_0^{r_0} pt^{p-1} |\{|Mf| > t\}| \, \mathrm{d}t \leq \int_0^{r_0} pt^{p-1} \frac{C}{t} \, \mathrm{d}t = pC \int_0^{r_0} t^{q-2} \, \mathrm{d}t < \infty$$

and thus, $Mf \in L^p(\mathbb{R}^d)$. Using this,

$$\begin{aligned}
||Mf||_p^p &= p \int_0^\infty t^{p-1} |\{Mf > t\}| \, \mathrm{d}t \\
&\leq p \int_0^\infty t^{p-1} \cdot \frac{3^d}{t} \int_{\{Mf > t\}} |f(x)| \, \mathrm{d}x \, \mathrm{d}t && \text{(Theorem 3.2.4)} \\
&= 3^d p \int_{\mathbb{R}^d} \int_0^{Mf(x)} t^{p-2} |f(x)| \, \mathrm{d}t \, \mathrm{d}x && \text{(Fubini's Theorem)} \\
&= \frac{3^d p}{p-1} \int_{\mathbb{R}^d} Mf(x)^{p-1} |f(x)| \, \mathrm{d}x \\
&\leq \frac{3^d p}{p-1} ||Mf||_p^{p-1} ||f||_p. && \text{(Hölder's Inequality)}
\end{aligned}$$

Dividing by $||Mf||_p^{p-1} < \infty$ gives $||Mf||_p \leq 3^d p' ||f||_p$.

If $f \in L^p(\mathbb{R}^d)$, then define $f_n := \mathbb{1}_{[-n,n]^d} \mathbb{1}_{\{|f| \leq n\}} f$. This ensures $f_n \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ and $|f_n| \uparrow |f|$. Thus for all $x \in \mathbb{R}^d$,

$$Mf(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| \, \mathrm{d}y = \sup_{Q \ni x} \lim_{n \to \infty} \frac{1}{|Q|} \int_Q |f_n(y)| \, \mathrm{d}y \leq \liminf_{n \to \infty} Mf_n(x)$$

by the Monotone Convergence Theorem. Finally, from Fatou's Lemma and the estimate obtained earlier,

$$||Mf||_p \leq ||\liminf_{n \to \infty} Mf_n||_p \leq \liminf_{n \to \infty} ||Mf_n||_p \leq 3^d p' \liminf_{n \to \infty} ||f_n||_p \leq 3^d p' ||f||_p. \qquad \square$$

3.3. Lebesgue Differentiation Theorem

The Lebesgue Differentiation Theorem (LDT) is the last bit of theory we need to cover before we can start discussing the Jessen–Marcinkiewicz–Zygmund Theorem. From the statement of the LDT, it is immediately clear why we needed to introduce the Hardy–Littlewood maximal operator and its weak L^1–boundedness. To improve the readability of the LDT and its proof, we first define the *Lebesgue points* of a function. We denote the cube with side lengths r and centre x by $Q_r(x)$.

Definition 3.3.1 (Lebesgue Point). *For a function* $f : \mathbb{R}^d \to \mathbb{C}$ *in* $L^1_{\text{loc}}(\mathbb{R}^d)$*, define* $\omega_f : \mathbb{R}^d \to \mathbb{R}$ *by*

$$\omega_f(x) := \limsup_{r\downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(y) - f(x)| \,\mathrm{d}y, \tag{3.1}$$

Then the points $x \in \mathbb{R}^d$ *for which* $\omega_f(x) = 0$ *are the Lebesgue points of* f*. The set* $L_f := \{x \in \mathbb{R}^d : \omega_f(x) = 0\}$ *denotes the set of Lebesgue points of* f*.*

Remark 3.3.2. Recall that functions that are equal almost everywhere are identified in L^p–spaces, and thus also in L^1_{loc}. It should be noted that identified functions do not necessarily have the same Lebesgue points: whether or not a point is a Lebesgue point depends on the specific version chosen for f.

Theorem 3.3.3 (Lebesgue Differentiation Theorem (LDT)). *Let* $f : \mathbb{R}^d \to \mathbb{C}$ *be a function such that* $f \in L^1_{\text{loc}}(\mathbb{R}^d)$*, then almost every* $x \in \mathbb{R}^d$ *is a Lebesgue point of* f*. That means* $\left|\mathbb{R}^d \setminus L_f\right| = 0$*, and*

$$\lim_{r\downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(y) - f(x)| \,\mathrm{d}y = 0, \quad x \in L_f. \tag{3.2}$$

In particular,

$$\lim_{r\downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} f(y)dy = f(x), \text{ and } |f(x)| \leq Mf(x) \text{ for all } x \in L_f. \tag{3.3}$$

Proof. The proof is based on the boundedness of the Hardy–Littlewood maximal operator, Theorem 3.2.4. Since $\omega_f \geq 0$ for all $x \in \mathbb{R}^d$, we need to prove that $|\{\omega_f > 0\}| = 0$. Define the cubes $I_k := (-k, k)^d$ and note that $|\{\omega_f > 0\}| = \lim_{k\to\infty} |\{\omega_f > 0\} \cap I_k|$ and for each $k \geq 1$,

$$\{\omega_f > 0\} \cap I_k = \{\omega_{\mathbb{1}_{I_k} f} > 0\} \cap I_k \subseteq \{\omega_{\mathbb{1}_{I_k} f} > 0\}.$$

It follows that it is enough to prove that $\{\omega_{\mathbb{1}_{I_k} f} > 0\} = 0$. Since $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and the cubes I_k are open, we can suppose that $f \in L^1(\mathbb{R}^d)$ when considering $\{\omega_{\mathbb{1}_{I_k} f} > 0\}$.

Let $\varepsilon > 0$ be arbitrary. Since $C_c(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, there exists a function $g \in C_c(\mathbb{R}^d)$ such that $||f - g||_1 < \varepsilon$. Hence, $\omega_g \equiv 0$ and by the triangle inequality

$$\omega_F \leq \omega_{f-g} + \omega_g \leq |f - g| + M(f - g).$$

Take $t > 0$, then by Proposition 2.2.9 and Theorem 3.2.4

$$|\{\omega_f > t\}| \leq |\{|f - g| > t/2\}| + |\{M(f - g) > t/2\}| \leq \frac{2}{t}||f - g||_1 + \frac{2 \cdot 3^d}{t}||f - g||_1 \leq \frac{2(1 + 3^d)}{t}\varepsilon,$$

for all $t > 0$. Since ε is arbitrary, this implies that $|\{\omega_f > t\}| = 0$. Now we find $|\{\omega_f > 0\}| = \lim_{n\to\infty} |\{\omega_f > 1/n\}| = 0$.

Finally, to prove (3.3), we note that

$$\begin{aligned}\left|f(x) - \frac{1}{|Q_r(x)|} \int_{Q_r(x)} f(y) \,\mathrm{d}y\right| &= \frac{1}{|Q_r(x)|} \left| \int_{Q_r(x)} f(x) - f(y) \,\mathrm{d}y\right| \\ &\leq \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(x) - f(y)| \,\mathrm{d}y.\end{aligned}$$

Since x is a Lebesgue point,

$$\limsup_{r\downarrow 0} \left|f(x) - \frac{1}{|Q_r(x)|} \int_{Q_r(x)} f(y) \,\mathrm{d}y\right| \leq \limsup_{r\downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |f(x) - f(y)| \,\mathrm{d}y = 0$$

and the first part of (3.3) follows. Also, $\frac{1}{|Q_r(x)|}\int_{Q_r(x)}|f(y)|\,\mathrm{d}y \le Mf(x)$ for all $x \in \mathbb{R}^d$ by definition of Mf. Thus, replacing f with $|f|$ in the first part of (3.3), gives the inequality in the second part of (3.3). □

In words, the Lebesgue Differentiation Theorem (LDT) states that, for any function that is locally integrable (i.e. $f \in L^1_{\text{loc}}$), the average of that function over any open cube (or ball) centred at x converges to $f(x)$ for almost every x. This is a generalisation of the Fundamental Theorem of Calculus in the sense that it applies to locally integrable functions instead of only to continuous functions.

Chapter 4

Jessen–Marcinkiewicz–Zygmund Theorem

After covering all the necessary theory, we can now present the Jessen–Marcinkiewicz–Zygmund (JMZ) Theorem, which forms the core of this thesis. As mentioned in Section 3.3, the Lebesgue Differentiation Theorem (LDT) tells us that:

The average of a locally integrable function f over any open cube (or ball) centred at x converges to $f(x)$ as the radius shrinks to zero.

Jessen, Marcinkiewicz, and Zygmund extended this result in their paper "Note on the differentiability of multiple integrals" [4]. Their main finding was that the cubes or balls used in the LDT can be replaced by the much more general collection of axis-aligned rectangles. While a cube or a ball in $\mathbb{R}^d$ is determined by a single size parameter – its side length or radius – an axis–aligned rectangle I can have d independent edge lengths. Because of this added flexibility compared to cubes or balls, the conditions imposed on f must be stronger to ensure that averages over the rectangles still converge to f almost everywhere as the diameter of I shrinks to zero. Throughout this chapter, $I \subseteq S$ is an axis–aligned rectangle and $S \subseteq \mathbb{R}^d$ denotes a bounded measurable set. We consider the set S instead of all of $\mathbb{R}^d$ because we require the measures of S and the rectangles $I \subseteq S$ to be finite for many of the arguments we use. Moreover, the JMZ Theorem is a *local* statement, just like the LDT, so it is enough to prove it only for all bounded measurable sets.

4.1. Definitions in JMZ Theorem

We use the set $S \subseteq \mathbb{R}^d$ and the axis–aligned rectangles $I \subseteq S$. The diameter of I is denoted by $\delta(I) := \sup\{d(x_1, x_2) : x_1, x_2 \in I\}$. As in Chapter 3, the Lebesgue measure of a set $A \in \mathbb{R}^d$ is denoted by $|A|$. Any two sides of the rectangles I are assumed to have a finite ratio, i.e. the rectangles are of finite eccentricity. Throughout the text, we frequently use the following two definitions.

Definition 4.1.1. *We say that the integral of a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is strongly differentiable at a point $x \in I \subseteq \mathbb{R}^d$, if*

$$\lim_{\delta(I)\to 0} \frac{1}{|I|} \int_I f(y)\, \mathrm{d}y \tag{4.1}$$

exists and is finite. The limit in (4.1) is called the strong derivative of the integral of f at x.

Definition 4.1.2. *For a function $f \in L^1(S)$ and axis–aligned rectangles $I \subseteq S$ that contain the point x,*

$$M_S f(x) := \sup_{I \ni x} \frac{1}{|I|} \int_I |f(y)|\, \mathrm{d}y, \tag{4.2}$$

$$m_S f(x) := \limsup_{\substack{\delta(I)\to 0 \\ I \ni x}} \frac{1}{|I|} \int_I |f(y)|\, \mathrm{d}y. \tag{4.3}$$

Remark 4.1.3. The operator $M_S f$ is the Hardy–Littlewood maximal operator (3.2.1) taken over all axis–aligned rectangles instead of just axis–aligned cubes. Because of this, $M_S f$ is called the *strong* maximal operator; the subscript S indicates this.

4.2. Lemmas for JMZ Theorem

Before we can give the full proof of the JMZ Theorem, we need to establish a few intermediate results. As indicated in the Introduction and Chapter 3, the proof of the JMZ Theorem relies heavily on maximal operators and their properties. Those properties are given and proved in this section.

We start with a very useful result which follows from the L^p–bound for the Hardy–Littlewood operator Mf (Thm. 3.2.5). The result was first found by Hardy and Littlewood back in 1930 [10].

Lemma 4.2.1. *Let $f : S \to \mathbb{C}$ be a measurable function with $S \subseteq \mathbb{R}$. If $f \in L(\log^+ L)^1(S)$, then $M_S f \in L^1(S)$, and*

$$\int_S M_S f(y)\,\mathrm{d}y \le A \int_S |f(y)|(1 + \log^+ |f(y)|)\,\mathrm{d}y + B,$$

where A and B are absolute constants.

Proof. We may assume that $f \ge 0$ since $M_S f(x) = M_S|f|(x)$ and split f into dyadic intervals. For each $k = 1, 2, \ldots$, we define

$$E_k := \{2^{k-1} \le f \le 2^k\}, \quad f_k = f \cdot \mathbb{1}_{E_k}.$$

For $k = 0$ we define $E_0 := \{f < 1\}$ and $f_0 = f \cdot \mathbb{1}_{E_0}$. Then $f = \sum_{k=0}^{\infty} f_k$ and by subadditivity, $M_S f \le \sum_{k=0}^{\infty} M_S f_k$.

By Theorem 3.2.5 with $d = 1$,

$$\int_S M_S f_k(y)\,\mathrm{d}y \le \left(\int_S (M_S f_k(y))^p\,\mathrm{d}y\right)^{1/p} \le 3p'\left(\int_S (f_k(y))^p\,\mathrm{d}y\right)^{1/p} \le \frac{3p}{p-1} 2^k |E_k|^{1/p},$$

where p' is the Hölder conjugate of p. Set $p = 1 + \frac{1}{k+1}$ for $k = 0, 1, \ldots$, then $\frac{1}{p-1} = k + 1$ and thus

$$\int_S M_S f_k(y)\,\mathrm{d}y \le 3(k+2)2^k |E_k|^{(k+1)/(k+2)}.$$

Summing over k yields

$$\int_S M_S f(y)\,\mathrm{d}y \le C \sum_{k=0}^{\infty} (k+1)2^k |E_k|^{(k+1)/(k+2)} \ . \tag{4.4}$$

Since $2^{k-1} < f \le 2^k$ on E_k, we have $f(y)(1 + \log^+ f(y)) \ge 2^{k-1}(k-1)(1 + \log 2)$ on E_k. Thus

$$\int_S f(y)(1+\log^+ f(y))\,\mathrm{d}y = \sum_{k=0}^{\infty}\left(\int_{E_k} f(y)(1+\log^+ f(y))\,\mathrm{d}y\right) \ge \frac{1}{2}(1+\log 2)\sum_{k=0}^{\infty}(k-1)2^k|E_k|. \tag{4.5}$$

In order to be able to compare (4.4) and (4.5), we need to relate $|E_k|^{(k+1)/(k+2)}$ to $|E_k|$. We split the indices into two classes, one for which the measure of the set is large, $|E_k| \ge 3^{-k}$, and one for which the measure is small, $|E_k| < 3^{-k}$. For k where $|E_k| \ge 3^{-k}$ we have

$$|E_k|^{(k+1)/(k+2)} = |E_k||E_k|^{-1/(k+2)} \le 3\,|E_k|.$$

If $|E_k| < 3^{-k}$, then

$$(k+1)2^k|E_k|^{(k+1)/(k+2)} \le (k+1)2^k 3^{-k} \le \tilde{C}(2/3)^k.$$

The sum of $(k+1)2^k|E_k|^{(k+1)/(k+2)}$ over all k for which $|E_k| < 3^{-k}$ is finite and can be estimated by some constant.

Combining the two classes we obtain, with finite constants $\tilde{A}$ and $\tilde{B}$,

$$C\sum_{k=0}^{\infty}(k+1)2^k|E_k|^{(k+1)/(k+2)} \leq \tilde{A}\Big(\sum_{k=0}^{\infty}(k-1)2^k|E_k|\Big) + \tilde{B}\,.$$

We can now relate (4.4) and (4.5) to find

$$\begin{aligned}\int_S M_S f(y)\,\mathrm{d}y &\leq C\sum_{k=0}^{\infty}(k+1)2^k|E_k|^{(k+1)/(k+2)} \\ &\leq \tilde{A}\Big(\sum_{k=0}^{\infty}(k-1)2^k|E_k|\Big) + \tilde{B} \\ &\leq A\int_S f(y)(1+\log^+ f(y))\,\mathrm{d}y + B\,.\end{aligned}$$

□

The result holds for $S \subseteq \mathbb{R}^2$ and functions $f \in L(\log^+ L)^1(S)$. In order to generalise this, we first need to take a step back to $S \subseteq \mathbb{R}$ and prove a comparable result for functions $f \in L(\log^+ L)^r(S)$ with $r \in \mathbb{N}$.

Lemma 4.2.2. *Let $f : S \to \mathbb{C}$ be a measurable function with $S \subseteq \mathbb{R}$. If $f \in L(\log^+ L)^r(S)$, with $r \in \mathbb{N}$, then $M_S f \in L(\log^+ L)^{r-1}(S)$, and*

$$\int_S M_S f(y)(1+\log^+ M_S f(y))^{r-1}\,\mathrm{d}y \leq A_r\int_S |f(y)|(1+\log^+|f(y)|)^r\,\mathrm{d}y + B_r, \tag{4.6}$$

where A_r and B_r depend only on r.

Proof. We start this proof by stating that it is sufficient to prove the result for non–negative and non–increasing functions f. The details of this are beyond the scope of this thesis and can be found in [11, p. 291].

Let $\varphi(t) := t(1+\log^+ t)^{r-1}$ and assume f is non–negative and non–increasing. Then $\varphi(f(x)) = |f(x)|(1+\log^+|f(x)|)^r$ is a non–increasing function of x since φ is convex. We can write the left–hand side of (4.6) as

$$\int_S \varphi(M_S f(y))\,\mathrm{d}y = \int_S \varphi\Big(\frac{1}{y}\int_0^y f(t)\,\mathrm{d}t\Big)\,\mathrm{d}y \leq \int_S \frac{1}{y}\,\mathrm{d}y\int_0^y \varphi(f(t))\,\mathrm{d}t, \tag{4.7}$$

by applying Jensen's Inequality (Thm. 2.2.4).

We have that $|f|(1+\log^+|f|)^r$ is integrable and thus $\varphi(f(x))$ is integrable, since f is non–negative by assumption. We can apply Lemma 4.2.1 to $\varphi(f(x))$. This allows us to write, for the right–hand side of (4.7),

$$\begin{aligned}\int_S \frac{1}{y}\,\mathrm{d}y\int_0^y \varphi(f(t))\,\mathrm{d}t &= \int_S \frac{1}{y}\Big(\int_0^y \varphi(f(t))\,\mathrm{d}t\Big)\,\mathrm{d}y \\ &= \int_S M_S\varphi(f)\,\mathrm{d}y \\ &\leq A\int_S |\varphi(f)|(1+\log^+|\varphi(f)|)\,\mathrm{d}y + B.\end{aligned}$$

Using the definition of φ and the bound $\log^+(1+\log^+ f) \leq 1+\log^+ f$ we then expand and rewrite this as:

$$\begin{aligned}
A\int_S |\varphi(f)|(1+\log^+|\varphi(f)|)\,\mathrm{d}y + B &= A\int_S [f\ (1+\log^+ f)^{r-1}](1+\log^+(f\ [1+\log^+ f]^{r-1}))\,\mathrm{d}y + B \\
&= A\int_S [f\ (1+\log^+ f)^{r-1}](1+\log^+(f)+\log^+([1+\log^+ f]^{r-1}))\,\mathrm{d}y + B \\
&\leq A\int_S [f(1+\log^+ f)^{r}] + [f(1+\log^+ f)^{r-1}]\cdot(r-1)\,(1+\log^+ f)\,\mathrm{d}y + B \\
&= Ar\int_S [f(1+\log^+ f)^{r}]\,\mathrm{d}y + B.
\end{aligned}$$

Recalling that it sufficed to show that (4.6) holds for non–negative and non–increasing f, we indeed find that $\int_S M_S f(y)(1+\log^+ M_S f(y))^{r-1}\,\mathrm{d}y \leq A_r \int_S |f(y)|(1+\log^+|f(y)|)^r\,\mathrm{d}y + B_r$. $\square$

4.3. The Jessen–Marcinkiewicz–Zygmund Theorem

Finally, we reach the main result of [4], the Jessen–Marcinkiewicz–Zygmund Theorem. The theorem is a generalisation of the Lebesgue Differentiation Theorem. Instead of being restricted to taking the average of a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ over only cubes or balls, the JMZ Theorem takes averages over arbitrary axis–aligned rectangles with positive side lengths.

Theorem 4.3.1 (Jessen–Marcinkiewicz–Zygmund, 1935). *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^{d-1}(S)$, then the integral of $f(x)$ is strongly differentiable at almost every $x \in S$, and the derivative is equal to $f(x)$.*

4.4. Proof of the JMZ Theorem

We want to apply Lemma 4.2.2 but this lemma is only stated for measurable functions in $S \subseteq \mathbb{R}$ and we are now considering measurable functions in $S \subseteq \mathbb{R}^d$. We work around this by applying the lemma coordinate–wise. For $i = 0, 1, \ldots, d-1$ define the iterated one–dimensional Hardy–Littlewood maximal operators,

$$\begin{aligned}
F_0(y) &:= |f(y)| \\
F_1(y) &:= M[F_0(\cdot, x_2, \ldots, x_d)](x_1) \\
F_2(y) &:= M[F_1(x_1, \cdot, x_3, \ldots, x_d)](x_2) \\
&\vdots \\
F_{d-1}(y) &:= M[F_{d-2}(x_1, \ldots, x_{d-2}, \cdot, x_d)](x_{d-1}),
\end{aligned}$$

where the M operator is similar to the Hardy–Littlewood maximal operator in Definition 3.2.1 but instead of the cube Q we take a one–dimensional interval that contains x_i. Note that for all i: $F_i > 0$, $F_i < \infty$ a.e., and $F_i \geq F_{i-1}$.

Lemma 4.2.2 is applied successively to each coordinate until we reach an inequality containing F_{d-1}. Take $r = d-1$. Since we know that $f \in L(\log^+ L)^{d-1}(S)$, we can apply Lemma 4.2.2 and then integrate over all remaining variables:

$$\int_S F_1(1+\log^+ F_1)^{d-2}\,\mathrm{d}y \leq A_{d,1}\int_S |f|(1+\log^+|f|)^{d-1}\,\mathrm{d}y + B_{d,1}.$$

This implies that $|F_1|(1+\log^+|F_1|)^{d-2} = F_1(1+\log^+ F_1)^{d-2} \in L^1(S)$ so we can now apply Lemma 4.2.2 with $r_1 = d-2$ and again integrate over all remaining variables:

$$\int_S F_2(1+\log^+ F_2)^{d-3}\,\mathrm{d}y \le A_{d,2}\int_S |F_1|(1+\log^+|F_1|)^{d-2}\,\mathrm{d}y + B_{d,2}.$$

Now we know that $|F_2|(1+\log^+|F_2|)^{d-3} \in L^1(S)$. The process is performed a total of $(d-1)$ times such that in the end we have an exponent of 0 on the left–hand side:

$$\int_S F_{d-1}(1+\log^+ F_{d-1})^{d-d}\,\mathrm{d}y \le A_{d,d-1}\int_S |F_{d-2}|(1+\log^+|F_{d-2}|)^{1}\,\mathrm{d}y + B_{d,d-1}.$$

Combining all the inequalities results in the relation

$$\int_S F_{d-1}\,\mathrm{d}y \le A_d\int_S |f|(1+\log^+|f|)^{d-1}\,\mathrm{d}y + B_d, \tag{4.8}$$

where A_d and B_d are constants that only depend on d.

Let $I := I_1\times I_2\times\ldots\times I_d \subseteq S$ be an arbitrary axis–aligned rectangle that contains the point $u=(u_1,u_2,\ldots,u_d)$. We can rewrite the d–multiple integral as an iterated integral by Fubini's Theorem (Thm. 2.1.15),

$$\begin{aligned}\frac{1}{|I|}\int_I |f(y)|\,\mathrm{d}y &= \frac{1}{|I_2|\cdots|I_d|}\int_{I_d}\cdots\int_{I_2}\left(\frac{1}{|I_1|}\int_{I_1}|f(t,x_2,\ldots,x_d)|\,\mathrm{d}t\right)\mathrm{d}x_2\cdots\mathrm{d}x_d\\ &\le \frac{1}{|I_3|\cdots|I_d|}\int_{I_d}\cdots\int_{I_3}\left(\frac{1}{|I_2|}\int_{I_2}|F_1(u_1,t,\ldots,x_d)|\,\mathrm{d}t\right)\mathrm{d}x_3\cdots\mathrm{d}x_d\\ &\ \vdots\\ &\le \frac{1}{|I_d|}\int_{I_d}F_{d-1}(u_1,\ldots,u_{d-1},x_d)\,\mathrm{d}x_d.\end{aligned}$$

Since this holds for arbitrary I,

$$\begin{aligned} m_S f(u) &= \limsup_{\substack{\delta(I)\to 0\\ I\ni u}} \frac{1}{|I|}\int_I |f(y)|\,\mathrm{d}y\\ &\le \limsup_{\substack{\delta(I_d)\to 0\\ I_d\ni x_d}} \frac{1}{|I_d|}\int_{I_d}F_{d-1}(u_1,\ldots,u_{d-1},x_d)\,\mathrm{d}x_d\\ &= F_{d-1}(u) \qquad\qquad \text{for a.e. } P\in S,\end{aligned} \tag{4.9}$$

where the last equality is due to the one–dimensional Lebesgue Differentiation Theorem (Thm. 3.3.3). Combining (4.8) and (4.9):

$$\int_S m_S f(y)\,\mathrm{d}y \le \int_S F_{d-1}(y)\,\mathrm{d}y \le A_d\int_S |f(y)|(1+\log^+|f(y)|)^{d-1}\,\mathrm{d}y + B_d, \tag{4.10}$$

for all $f\in L(\log^+ L)^{d-1}(S)$. Apply (4.8) to the function λf, where $\lambda>0$ is a constant, to find

$$\int_S \lambda m_S f(y)\,\mathrm{d}y \le \int_S \lambda F_{d-1}\,\mathrm{d}y \le A_d\int_S |\lambda f|(1+\log^+|\lambda f|)^{d-1}\,\mathrm{d}y + B_d. \tag{4.11}$$

Let $\varepsilon>0$ be arbitrary and choose λ so large that $B_d/\lambda < \frac{1}{2}\varepsilon$. By Lemma 2.2.6, f can be decomposed as $f(y)=\varphi(y)+\psi(y)$ where φ is continuous and $\psi := f-\varphi$ satisfies

$$\int_S |\psi(y)|\,\mathrm{d}y < \varepsilon, \tag{4.12}$$

$$A_d\int_S |\psi(y)|(1+\log^+|\lambda\psi(y)|)^{d-1}\,\mathrm{d}y + \frac{B_d}{\lambda} < \varepsilon. \tag{4.13}$$

Note that (4.13) can be satisfied as a result of Lemma 2.2.6 since we defined $L(\log^+ L)^{d-1}(S)$ as the space of functions g for which $\int_S |g|(1+\log^+|g|)^{d-1}\,\mathrm{d}y < \infty$, i.e. those functions g for which $|g|(1+\log^+|g|)^{d-1} \in L^1(S)$.

We apply (4.10) to $\lambda\psi$ and divide by λ,

$$\int_S m_S\psi(y)\,\mathrm{d}y \leq A_d \int_S |\psi(y)|(1+\log^+|\lambda\psi(y)|)^{d-1}\,\mathrm{d}y + \frac{B_d}{\lambda} < \varepsilon. \tag{4.14}$$

For the last steps of the proof, we need to define two sets, $E_1(\varepsilon) := \{y \in S : |\psi(y)| > \sqrt{\varepsilon}\}$ and $E_2(\varepsilon) := \{y \in S : m_S\psi(y) > \sqrt{\varepsilon}\}$, and denote their union by $E(\varepsilon)$. We apply Chebyshev (Prop. 2.2.9) with $p = 1$ to these sets. Recall that $||f||_{p,\infty} = \sup_{t>0}\ t|\{|f| > t\}|^{1/p}$. Then,

$$|E_1(\varepsilon)| = |\{|\psi(y)| > \sqrt{\varepsilon}\}| \leq \frac{1}{\sqrt{\varepsilon}} \int_S |\psi(y)|\,\mathrm{d}y < \frac{\varepsilon}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}.$$

Similarly for $E_2(\varepsilon)$,

$$|E_2(\varepsilon)| = |\{m_S\psi(y) > \sqrt{\varepsilon}\}| \leq \frac{1}{\sqrt{\varepsilon}} \int_S m_S\psi(y)\,\mathrm{d}y < \frac{\varepsilon}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}.$$

Thus, $|E(\varepsilon)| \leq |E_1(\varepsilon)| + |E_2(\varepsilon)| < 2\sqrt{\varepsilon}$.

With the decomposition of f, and $I \ni x$ any axis–aligned rectangle, we write,

$$\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x) = \left[\frac{1}{|I|}\int_I \varphi(y)\,\mathrm{d}y - \varphi(x)\right] + \left[\frac{1}{|I|}\int_I \psi(y)\,\mathrm{d}y - \psi(x)\right].$$

For $n \in \mathbb{N}$ set $\varepsilon_n := 4^{-n}$ and $E_n := E(\varepsilon_n)$. Then

$$|E_n| \leq 2\sqrt{\varepsilon_n} = 2 \cdot 2^{-n},$$

and thus $\sum_{n=1}^{\infty} |E_n| < \infty$. Define

$$S_0 := S \setminus \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} E_n = S \setminus \limsup_{n\to\infty} E_n\,.$$

The Borel–Cantelli Lemma tells us that

$$|\limsup_{n\to\infty} E_n| = 0\,,$$

so S_0 has full measure.

By definition of $\limsup E_n$, if there exists an $x \notin \limsup E_n$, then we can choose an $N = N(x)$ such that $x \notin E_n$ for all $n \geq N$. Let N be such an index. Because $x \notin E_n$ for all $n \geq N$,

$$\limsup_{\delta(I)\to 0} \left|\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x)\right| \leq m_S\psi(x) + |\psi(x)| \leq 2\sqrt{\varepsilon_n} = 2 \cdot 2^{-n}\,. \tag{4.15}$$

Recall that φ is continuous and therefore $\left(\frac{1}{|I|}\int_I \varphi(P)\,\mathrm{d}P - \varphi(P_0)\right) \to 0$ as I shrinks to x. Since the right–hand side of (4.15) tends to zero as $n \to \infty$, we find that

$$\limsup_{\delta(I)\to 0} \left|\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x)\right| = 0 \qquad \text{for every} \quad x \in S_0\,.$$

In words: the integral of f is strongly differentiable on a set of full measure, i.e. almost everywhere, and the derivative is equal to f. □

4.5. Sharpness

The restriction in the condition $f \in L(\log^+ L)^{d-1}(S)$ for $S \subseteq \mathbb{R}^d$ is necessary and cannot be weakened without the theorem failing for some f. We can formally prove this sharpness using the lemma stated below.

First, let $d \geq 2$ and take a measurable and increasing function $\varphi : [0,\infty) \to [0,\infty)$ such that

$$\varphi(0) = 0, \qquad \liminf_{t\to\infty} \frac{\varphi(t)}{t} > 0. \tag{4.16}$$

This limit inferior indicates that as $t \to \infty$, φ grows at least linearly. Define the class of measurable functions,

$$L^{\varphi} := \{f : S \to \mathbb{R} \; : \int_S \varphi(|f|)\,\mathrm{d}x < \infty\}.$$

The Jessen–Marcinkiewicz–Zygmund Theorem tells us that the integrals of functions $f \in L^{\varphi}(S)$ with $\varphi(t) := t(1 + \log^+ t)^{d-1}$ are strongly differentiable.

Lemma 4.5.1. *Let φ be as defined in (4.16) and let $L^{\varphi}(S)$ be the corresponding space of measurable functions. For a bounded measurable set $E \subseteq S$ and a parameter $0 < \alpha < 1$ define*

$$\sigma_\alpha(E) := \{x \in S \,:\, I \ni x,\, |E \cap I| > \alpha|I|\},$$

where $I \subseteq S$ are axis–aligned rectangles. If the Jessen–Marcinkiewicz–Zygmund Theorem (Thm. 4.3.1) holds for every $f \in L^{\varphi}(S)$, then

$$|\sigma_\alpha(E)| \leq C_{(d,\varphi)}\varphi\left(\frac{1}{\alpha}\right)|E| \tag{4.17}$$

hold for all E and all α. The constant C only depends on d and φ.

Proof. We prove the lemma by assuming that inequality (4.17) is false and then show that there exists a function $f \in L^{\varphi}(S)$ for which the JMZ Theorem fails.

Let constants $C_n > 0$ be chosen such that

$$\sum_{n\geq 1} \frac{1}{C_n} < \frac{1}{2}\varphi(1).$$

By assumption, for every $n \in \mathbb{N}$ there exists a bounded and measurable set E_n and a constant $\alpha_n \in (0,1)$ such that

$$|\sigma_{\alpha_n}(E_n)| > C_n\varphi\left(\frac{1}{\alpha_n}\right)|E_n|. \tag{4.18}$$

Because of the properties of the Lebesgue measure, both sides of (4.18) scale in the same way. Rescale and/or translate every E_n such that

$$E_n \subseteq Q_0 := [0,1]^d, \quad \delta(E_n) < 2^{-n}.$$

Fix an n. Let $\mathcal{D}_n$ be the collection of all dyadic cubes with side lengths 2^{-n} in Q_0. Because $\delta(E_n) < 2^{-n}$, each cube $Q \in \mathcal{D}_n$, Q and E_n are either disjoint or $E_n \subseteq Q$. Set

$$\mathcal{R}_n := \{Q \in \mathcal{D}_n \,:\, |E_n| > \alpha_n|Q|\}.$$

All $Q \in \mathcal{R}_n$ are pairwise disjoint, satisfy

$$\bigcup_{Q\in\mathcal{R}_n} Q = \sigma_{\alpha_n}(E_n),$$

and thus the union has measure $|\sigma_{\alpha_n}(E_n)|$ by construction. For each $Q \in \mathcal{D}_n$ let E_n^Q be the translated version of E_n such that it is contained in Q. Set $F_n := \bigcup_{Q\in\mathcal{D}_n} E_n^Q$ and define the simple function

$$f_n := \tfrac{1}{\alpha_n}\mathbb{1}_{F_n}\,.$$

Set $f := \sup_{n\geq 1} f_n$. Then, using (4.18) and the choice for constants C_n,

$$\begin{aligned}\int_{Q_0}\varphi(f)\,\mathrm{d}y &\leq \sum_{n=1}^{\infty}\int_{Q_0}\varphi(f_n)\,\mathrm{d}y\\ &=\sum_{n=1}^{\infty}\varphi\left(\frac{1}{\alpha_n}\right)|F_n|\\ &=\sum_{n=1}^{\infty}2^{nd}|E_n|\varphi(\tfrac{1}{\alpha_n})\\ &<\sum_{n=1}^{\infty}\frac{2^{nd}}{C_n}|\sigma_{\alpha_n}(E_n)|\\ &\leq |Q_0|\sum_{n=1}^{\infty}\frac{1}{C_n}<\frac{1}{2}\varphi(1)\,.\end{aligned}\tag{4.19}$$

Hence $f \in L^{\varphi}(Q_0)$.

Fix n and a point $x \in Q_0$. Because the dyadic cubes tile Q_0, x lies in some $Q \in \mathcal{D}_n$, and thus in the translated and scaled set E_n^Q. Inside that same cube Q there is a subcube $R \subseteq Q$ for which $|E_n^Q \cap R| > \alpha_n|R|$. Therefore,

$$\frac{1}{|R|}\int_R f(y)\mathrm{d}y \geq \frac{1}{|R|}\int_R f_n(y)\mathrm{d}y = \frac{1}{\alpha_n|R|}|E_n^Q\cap R| > 1\,.$$

Because $\delta(R) \leq \delta(E_n^Q) \leq 2^{-n} \to 0$ as $n \to \infty$, the Jessen–Marcinkiewicz–Zygmund Theorem would force $f(x) \geq 1$ for a.e. x. But then

$$\int_{Q_0}\varphi(f)\,\mathrm{d}y \geq \varphi(1)\,|Q_0| = \varphi(1)\,,$$

contradicting the bound found in (4.19). This contradiction implies that the assumption in (4.18) must be false. This proves the lemma. ☐

With Lemma 4.5.1 established, we can now prove sharpness of the condition $f \in L(\log^+ L)^{d-1}(S)$ in the JMZ Theorem. The following theorem formally states this sharpness.

Theorem 4.5.2. *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If for every $f \in L^{\varphi}(S)$, where φ is as in (4.16), the integral of f is strongly differentiable almost everywhere, then $\varphi(t) > ct(1+\log^+ t)^{d-1}$ for some $c > 0$. In other words, $f \in L(\log^+ L)^{d-1}(S)$.*

Proof. We prove the theorem for the unit cube Q. Note that this is sufficient since, due to properties of the Lebesgue measure, Lemma 4.5.1 is scale– and translation–invariant, so the inequality automatically propagates to every translate and every dilate of the unit cube, and therefore to any bounded measurable set $S \subseteq \mathbb{R}^d$.

Let $Q = [0,1]^d$, $|Q|= 1$, fix $\alpha \in (0,1)$. For every $u = (u_1,\ldots,u_{d-1},u_d)$ with $u_j \geq 1$ and $u_1\cdots u_{d-1} \leq \frac{1}{\alpha}$ for $j = 1,\ldots,d-1$ and $0 \leq u_d \leq \frac{1}{\alpha u_1\cdots u_{d-1}}$. Denote the set of all these $u \in \mathbb{R}^d$ by S_α.

For $u \in S_\alpha$ define the axis–aligned rectangle

$$I(u) := (0, u_1] \times \cdots \times (0, u_{d-1}] \times (0, u_d]\,,$$

then $|I(u)| = u_1 \cdots u_{d-1} u_d \le \frac{1}{\alpha}$. Because $u_i \ge 1$ for $j = 1, \ldots, d-1$ while $0 \le u_d \le 1$, the slice of $I(u)$ that sits inside the unit cube Q is $J(u) := [0,1]^{d-1} \times [0, u_d]$ with $|J(u)| = u_d$. Hence,

$$\frac{|E \cap I(u)|}{|I(u)|} = \frac{u_d}{u_1 \cdots u_{d-1} u_d} = \frac{1}{u_1 \cdots u_{d-1}} \ge \alpha\,.$$

Therefore, $u \in \sigma_\alpha(Q)$ as in Lemma 4.5.1, so

$$S_\alpha \subseteq \sigma_\alpha(Q)\,. \tag{4.20}$$

Let $S_\alpha(u_1, \ldots, u_{d-1}) := \{(u_1, \ldots, u_{d-1}) \,:\, u_i \in S_\alpha,\;\; i = 1, \ldots, d-1\}$ Since $u_d \in [0, 1/(\alpha u_1 \cdots u_{d-1})]$, Fubini (Thm. 2.1.15) yields

$$\begin{aligned} |S_\alpha| &= \int_{S_\alpha(u_1,\ldots,u_{d-1})} \frac{1}{\alpha u_1 \cdots u_{d-1}}\, \mathrm{d}u_1 \cdots \mathrm{d}u_{d-1} \\ &= \frac{1}{\alpha} \int_{S_\alpha(u_1,\ldots,u_{d-1})} \frac{1}{u_1 \cdots u_{d-1}}\, \mathrm{d}u_1 \cdots \mathrm{d}u_{d-1}\,. \end{aligned}$$

Set $v_i := \log u_i$ for all $i = 1, \ldots, d-1$. Then $u_i = e^{v_i}$, with $v_i \ge 0$ and $\mathrm{d}u_i = e^{v_i}\mathrm{d}v_i$. Thus,

$$\frac{1}{u_1 \cdots u_{d-1}}\, \mathrm{d}u_1 \cdots \mathrm{d}u_{d-1} = \mathrm{d}v_1 \cdots \mathrm{d}v_{d-1}\,.$$

The inequality $u_1 \cdots u_{d-1} \le 1/\alpha$ becomes $v_1 + \cdots + v_{d-1} \le \log \frac{1}{\alpha}$. Hence

$$|S_\alpha| = \frac{1}{\alpha} \int_{\substack{v_i \ge 0 \\ \sum v_i \le \beta}} \mathrm{d}v_1 \cdots \mathrm{d}v_{d-1} = \frac{1}{\alpha} \frac{(\log \frac{1}{\alpha})^{d-1}}{(d-1)!}\,.$$

Combining this with (4.20) gives the lower bound for $\sigma_\alpha(Q)$:

$$|\sigma_\alpha(Q)| \ge \frac{1}{\alpha} \frac{(\log \frac{1}{\alpha})^{d-1}}{(d-1)!}\,, \quad 0 < \alpha \le 1\,.$$

It follows from Lemma 4.5.1 that

$$\frac{1}{\alpha} \frac{(\log \frac{1}{\alpha})^{d-1}}{(d-1)!} \le |\sigma_\alpha(Q)| \le C_{d,\varphi}\, \varphi\Big(\frac{1}{\alpha}\Big) |Q|\,.$$

Setting $t = \frac{1}{\alpha}$ yields

$$\varphi(t) \ge \frac{1}{(d-1)!\;\, C_{d,\varphi}} t(1 + \log t)^{d-1} = ct(1 + \log t)^{d-1}\,, \quad t > 1\,. \qquad \square$$

The Jessen–Marcinkiewicz–Zygmund Theorem proved that the condition $f \in L(\log^+ L)^{d-1}$ is sufficient for strong differentiability of the integral of f. Theorem 4.5.2 proves that $f \in L(\log^+ L)^{d-1}$ is also a necessary condition. Together, these results show that the condition is optimal: it is sharp.

Chapter 5

Zygmund's Conjecture

Thirty years after the publication of the original article by Jessen, Marcinkiewicz, and Zygmund [4], Antoni Zygmund published an article in which he extended the original result [5]. He showed that if we assume k size parameters for the rectangles I in the Jessen–Marcinkiewicz–Zygmund (JMZ) Theorem, the theorem holds for $f \in L(\log^+ L)^{k-1}(S)$. From this, he then conjectured that if the rectangles were determined by positive functions depending on k variables, the theorem would also hold for $f \in L(\log^+ L)^{k-1}(S)$. This was shown to be false by F. Soria in [12]. In this section, we present the extended result by Zygmund and discuss his conjecture and the counterexamples that disprove it.

Throughout this chapter, we use the familiar conditions on $S \subseteq \mathbb{R}^d$ and $I \subseteq S$. This means that S is again a bounded measurable set and that the rectangles $I \subseteq S$ are axis–aligned.

5.1. Generalisation by Zygmund

In 1967, Antoni Zygmund published a postscript, [5], to [4] in which he extended the JMZ Theorem. To prove this, we need three lemmas as intermediate results. Throughout this section, unless stated otherwise, we use $p \in (1, \infty)$. We consider axis–aligned rectangles I with at most k *different side lengths*, where $1 \leq k \leq d$. The operators $M_S f(x)$ and $m_S f(x)$ are defined as in (4.2) and (4.3), respectively, but keep in mind that the rectangles are now more restricted.

When considering rectangles with $1 \leq k \leq d$ distinct side lengths, we can sharpen the factors in the inequalities in Theorem 3.2.5 and Lemma 4.2.1.

Lemma 5.1.1. *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L^p(S)$, and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths, then $M_S f \in L^p(S)$ and*

$$||M_S f||_p \leq \frac{A}{(p-1)^k} ||f||_p,$$

where A depends on d and p, but is bounded over any finite range of p.

Proof. We may assume that $f \geq 0$ and perform induction on k. Suppose that $k > 1$ and that the statement holds for $k-1$. Denote the k distinct side lengths of I by $h_1, h_2, \ldots, h_k$. We may assume that these sizes correspond to groups of coordinates $x_1, x_2, \ldots, x_d$.

Define $x = (x', x'')$ where x' denotes all coordinates corresponding to the side length h_1 and x'' denotes the remaining coordinates. Set

$$S' := \{x' : (x', x'') \in S\}$$
$$S'' := \{x'' : (x', x'') \in S\}$$

Similarly, we write $I = Q' \times I''$ where Q' is a cube in the space of x' and I'' is a rectangle in the space of x''. Note that I'' has $(k-1)$ different size parameters, not k. The condition $I \ni x$

implies that $x' \in Q'$ and $x'' \in I''$ and if we set

$$f^*(x', x'') := \sup_{I'' \ni x''} \frac{1}{|I''|} \int_{I''} f(x', y'') \, \mathrm{d}y'', \tag{5.1}$$

$$f^{**}(x', x'') := \sup_{Q' \ni x'} \frac{1}{|Q'|} \int_{Q'} f^*(y', x'') \, \mathrm{d}y', \tag{5.2}$$

then from

$$\frac{1}{|I|} \int_I f(y) \, \mathrm{d}y = \frac{1}{|Q'|} \int_{Q'} \left(\frac{1}{|I''|} \int_{I''} f(y', y'') \, \mathrm{d}y'' \right) \mathrm{d}y' \,,$$

we find that

$$M_S f(x) \leq f^{**}(x', x'').$$

Hence, by applying Theorem 3.2.5 and then Lemma 5.1.1 with $(k-1)$, we obtain using Fubini (Thm 2.1.15),

$$\begin{aligned}
\int_S (M_S f(x))^p \, \mathrm{d}x &\leq \int_{S''} \left(\int_{S'} [f^*(x', x'')]^p \, \mathrm{d}x' \right) \mathrm{d}x'' \\
&\leq \left(\frac{3p}{p-1} \right)^p \int_{S''} \left(\int_{S'} [f^*(x', x'')]^p \, \mathrm{d}x' \right) \mathrm{d}x'' \\
&= \left(\frac{3p}{p-1} \right)^p \int_{S'} \left(\int_{S''} [f^*(x', x'')]^p \, \mathrm{d}x'' \right) \mathrm{d}x' \\
&\leq \left(\frac{3p}{p-1} \right)^p \left(\frac{\tilde{A}}{p-1} \right)^{p(k-1)} \int_{S'} \left(\int_{S''} [f(x', x'')]^p \, \mathrm{d}x'' \right) \mathrm{d}x' \\
&= \left(\frac{A}{(p-1)^k} \right)^p ||f||_p^p.
\end{aligned}$$

Taking powers of $1/p$ gives the desired result. ☐

With the analogy for Theorem 3.2.5, we can prove our next result. It follows from Lemma 5.1.1 and, as expected from the statement, the proof is analogous to the proof of Lemma 4.2.1. Because the proof is just a repetition of steps already performed when proving Lemma 4.2.1, we do not state it explicitly and refer the reader to Appendix A.1.

Lemma 5.1.2. *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^k(S)$ and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths. Then $M_S f \in L^1(S)$ and*

$$\int_S M_S f(y) \, \mathrm{d}y \leq A \int_S |f(y)| (1 + \log^+ |f(y)|)^k \, \mathrm{d}y + B,$$

where A and B depend only on d.

Using Lemma 5.1.2 we can now prove an upper bound for $m_S f(x)$ when $f \in L(\log^+ L)^{k-1}(S)$. We need it when proving Zygmund's slightly generalised version of the Jessen–Marcinkiewicz–Zygmund Theorem because we need to eventually bound a limit. In the proof of the JMZ Theorem, this lemma would roughly correspond to the steps taken up to and including (4.10).

Lemma 5.1.3. *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^{k-1}(S)$ and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths. Then $m_S f \in L^1(S)$ and*

$$\int_S m_S f(y) \, \mathrm{d}y \leq A \int_S |f(y)| (1 + \log^+ |f(y)|)^{k-1} \, \mathrm{d}y + B \, |S'|,$$

where A and B depend only on d.

Proof. The initial setup is the same as in the proof of Lemma 5.1.1. We also again set

$$f^*(x', x'') := \sup_{I'' \ni x''} \frac{1}{|I''|} \int_{I''} f(x', y'') \,\mathrm{d}y''.$$

We may suppose that $f \geq 0$. Under the hypotheses of the lemma, $f(x) = f(x', x'')$, as a function of x'', is in $L(\log^+ L)^{k-1}(S'')$ for almost every $x' \in S'$. For such a fixed x' consider the function $f^*(x', x'')$. By Lemma 5.1.2, with $(k-1)$,

$$\int_{S''} f^*(x', x'') \,\mathrm{d}x'' \leq A \int_{S''} f(x', x'')(1 + \log^+ f(x', x''))^{k-1} \,\mathrm{d}x'' + B.$$

Integrating over $x' \in S'$ gives

$$\int_{S' \times S''} f^*(x', x'') \,\mathrm{d}x' \,\mathrm{d}x'' = \int_S f^*(x', x'') \,\mathrm{d}x' \,\mathrm{d}x'' \leq A \int_S f(x)(1 + \log^+ f(x))^{k-1} \,\mathrm{d}x + B\,|S'|. \quad (5.3)$$

It follows that for almost every $x'' \in S''$ the function $f^*(x', x'')$ is integrable in x' over S'. Fix such an x''. If $x = (x', x'') \subseteq I = Q' \times I''$, then since,

$$\frac{1}{|I|} \int_I f(y) \,\mathrm{d}y = \frac{1}{|Q'|} \int_{Q'} \left(\frac{1}{|I''|} \int_{I''} f(y', y'') \,\mathrm{d}y'' \right) \mathrm{d}y',$$

we find that

$$\frac{1}{|I|} \int_I M_S f(y) \,\mathrm{d}y \leq \frac{1}{|Q'|} \int_{Q'} f^*(y', x'') \mathrm{d}y'.$$

Thus, by the Lebesgue Differentiation Theorem (Thm. 3.3.3),

$$m_S f(x) \leq f^*(x', x'') \quad (5.4)$$

for almost all x'. It follows that (5.4) holds for almost all $x \in S$ and thus by (5.3),

$$\int_S m_S f(y) \,\mathrm{d}y \leq A \int_S f(x)(1 + \log^+ f(x))^{k-1} \,\mathrm{d}x + B\,|S'|\,.$$ □

These three lemmas naturally lead to Zygmund's extended version of the Jessen–Marcinkiewicz–Zygmund Theorem (Thm. 4.3.1). The proof of the theorem stated next follows exactly the same steps as the proof of the JMZ Theorem, starting from (4.11), if we replace Lemma 4.2.1 by Lemma 5.1.2. An explicit proof is given in Appendix A.1.

Theorem 5.1.4 (Zygmund, 1967). *Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^{k-1}(S)$ and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths, then the integral of $f(x)$ is strongly differentiable with respect to I at almost every $x \in S$, and the derivative is equal to $f(x)$.*

5.2. Zygmund's Conjecture

Based on the theorem he proved in 1967, Zygmund conjectured that if the rectangles over which the averages are taken are constructed from positive functions depending on $1 \leq k \leq d$ positive parameters, then the integral of $f \in L(\log^+ L)^{k-1}$ would still be strongly differentiable almost everywhere.

To simplify notation, we denote the d–dimensional rectangle with side lengths

$$u_1 \times u_2 \times \cdots \times u_d$$

as the bounded and measurable set

$$[u_1, u_2, \dots, u_d] \subseteq \mathbb{R}^d.$$

The collection of all d–dimensional rectangles with such side lengths is then defined as

$$\mathcal{B} := \{[u_1, \dots, u_d] \,:\, u_i \in \mathbb{R} \quad \text{for all} \quad 1 \leq i \leq d\}.$$

In the context of this thesis, a collection of bounded measurable sets is called a *basis*. So, $\mathcal{B}$ is a basis of d–dimensional rectangles.

Conjecture 5.2.1 (Zygmund's Conjecture). *Let $f \in L(\log^+ L)^{k-1}(S)$ with $S \subseteq \mathbb{R}^d$ and $1 \leq k \leq d$. Let there be d positive functions $\varphi_i : \mathbb{R}^k \to \mathbb{R}$ that are non–decreasing in each variable. Define the basis*

$$\mathcal{B} := \{[\varphi_1(t_1, \dots, t_k), \dots, \varphi_d(t_1, \dots, t_k)] \,:\, t_1, \dots, t_k > 0\}.$$

If $\varphi_1, \dots, \varphi_d$ can take arbitrarily small values, then for $I \in \mathcal{B}$

$$\lim_{\substack{\delta(I) \to 0 \\ I \ni x}} \frac{1}{|I|} \int_I |f(y)| \, \mathrm{d}y = f(x) \quad \text{for almost all} \quad x \in S\,.$$

Remark 5.2.2. If this convergence holds with $I \in \mathcal{B}$, we say that the basis $\mathcal{B}$ *differentiates* the space $L(\log^+ L)^{k-1}(S)$.

Clearly, if $k = d$, the conjecture is just the Jessen–Marcinkiewicz–Zygmund Theorem as stated in its original form (Thm. 4.3.1) so we are really only interested in if the conjecture holds for $1 \leq k < d$.

If, for every $i = 1, \dots, d$, we let $\varphi_i(t_1, \dots, t_k) = t_j$ for some $1 \leq j \leq k$, then by Zygmund's Theorem (Thm. 5.1.4) the conjecture holds. Antonio Córdoba presented the first non–trivial result in the direction of the conjecture [13]. In his article, he showed that the conjecture holds for the collection of rectangles $\mathcal{B} := \{[s, t, \varphi(s,t)] : s > 0, t > 0\}$.

5.3. Counterexamples by F. Soria

The first counterexamples to Zygmund's conjecture were constructed by F. Soria [12]. He starts by constructing counterexamples in $\mathbb{R}^3$ and later extends them to higher dimensions. The main idea of Soria's counterexample in $\mathbb{R}^3$ is constructing a basis $\mathcal{B}$ in the bounded measurable set $S \subseteq \mathbb{R}^3$ that uses only two parameters but behaves "as badly" as the whole collection of rectangles in S; the basis $\mathcal{B}_3 := \{[x, y, z] \in S\}$.

We start by introducing a variant on the Hardy–Littlewood maximal operator (3.2.1). For each basis $\mathcal{B}$ we define the maximal operator $M_{\mathcal{B}} f$.

Definition 5.3.1. *For a function $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and a basis $\mathcal{B}$ in $\mathbb{R}^d$, define $M_{\mathcal{B}} f : \mathbb{R}^d \to [0, \infty]$ by*

$$M_{\mathcal{B}} f(x) = \sup_{\substack{I \in \mathcal{B} \\ I \ni x}} \frac{1}{|I|} \int_I |f(y)| \, \mathrm{d}y,$$

where the supremum is taken over all rectangles $I \in \mathcal{B}$ that contain x.

The first counterexample is restricted to $\mathbb{R}^3$ and is then readily generalised to $\mathbb{R}^d$. We need the following result to start our construction.

Proposition 5.3.2. *Let $\mathcal{B}$ be a collection of rectangles in $\mathbb{R}^3$. Given $I \in \mathcal{B}$, denote the side lengths of I in the x, y, and z directions by 1_I, 2_I, and 3_I, respectively. Assume that there exists a fixed direction, say the x–direction, with the following property: If I and I' are elements of $\mathcal{B}$ and $1_I > 1_{I'}$, then either $2_I \geq 2_{I'}$ or $3_I \geq 3_{I'}$. In that case*

$$|\{M_{\mathcal{B}}f(x) > a\}| \leq C \int_{\mathbb{R}^3} \frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right) \mathrm{d}x\,, \tag{5.5}$$

where the constant C depends only on the dimension.

Proof. Write $x = (t, u)$ with $t \in \mathbb{R}$ and $u = (y, z) \in \mathbb{R}^2$. Let $\mathcal{B}_1$ be the collection of all intervals on the t–axis and define $M_{\mathcal{B}_1}f$ as in Definition 5.3.1. Let $\mathcal{R}$ be the collection of all axis–aligned rectangles in the u–plane and let $M_S f(u)$ be as in (4.2).

For a fixed point $(t, u) \in \mathbb{R}^3$, choose a rectangle $I \in \mathcal{B}$ that contains (t, u) and has side lengths that satisfy $1_I \geq 2_I \geq 3_I$. Project I onto the t–axis and the u–plane: the interval $A \subseteq \mathbb{R}$ of length 1_I through the point t, and a rectangle $B \subseteq \mathbb{R}^2$ with side lengths $2_I, 3_I$ through the point u. Hence $I = A \times B$ and $|I|= |A|\,|B|$. For fixed $s \in A$

$$\frac{1}{|B|}\int_B |f(s,v)|\mathrm{d}v \leq M_S f(s,u)$$

since B is one of the rectangles over which $M_S f$, as a function of u, takes its supremum. By Fubini (Thm. 2.1.15) we write, for arbitrary $I \in \mathcal{B}$,

$$\begin{aligned}\frac{1}{|I|}\int_I |f(x)|\,\mathrm{d}x &= \frac{1}{|A|}\int_A \left(\frac{1}{|B|}\int_B |f(s,v)|\,\mathrm{d}v\right)\mathrm{d}s\\ &\leq \frac{1}{|A|}\int_A M_S f(s,u)\,\mathrm{d}s\\ &= M_{\mathcal{B}_1}[M_S f(\,\cdot\,,u)](t)\,.\end{aligned}$$

Taking the supremum over $I \ni (t,u) = x$

$$M_{\mathcal{B}}f(x) \leq M_{\mathcal{B}_1}[M_S f(\,\cdot\,,u)](t) \quad \text{for a.e. } (t,u) \in \mathbb{R}^3. \tag{5.6}$$

For $\alpha > 0$ set $E_\alpha := \{M_{\mathcal{B}}f > \alpha\}$. Applying (5.6) and Theorem 3.2.4:

$$|E_\alpha| \leq |\{(t,u)\,:\, M_{\mathcal{B}_1}[M_{\mathcal{R}}f(\,\cdot\,,u)](t) > \tfrac{\alpha}{C_0}\}| \leq \frac{C_1}{\alpha}\int_{\mathbb{R}^2}\int_{\mathbb{R}} M_{\mathcal{R}}f(t,u)\,\mathrm{d}t\,\mathrm{d}u\,. \tag{5.7}$$

By Lemma 4.2.1,

$$\int_{\mathbb{R}} M_{\mathcal{R}}f(t,u)\,\mathrm{d}t \leq C_2\int_{\mathbb{R}} |f(t,u)|(1 + \log^+|f(t,u)|)\,\mathrm{d}t \quad \text{for a.e. } u \in \mathbb{R}^2,$$

where we take C_2 so large that B is irrelevant. Integrating this bound in u yields, by Fubini (Thm. 2.1.15),

$$\int_{\mathbb{R}^2}\int_{\mathbb{R}} M_{\mathcal{R}}f(t,u)\,\mathrm{d}t\,\mathrm{d}u = \int_{\mathbb{R}^3} M_{\mathcal{R}}f(t,u)\,\mathrm{d}(t,u) \leq C_2\int_{\mathbb{R}^3}|f(x)|(1 + \log^+|f(x)|)\,\mathrm{d}x\,. \tag{5.8}$$

For $\alpha > 0$ we have

$$\frac{|f|}{\alpha}(1 + \log^+|f|) = \frac{|f|}{\alpha}\left(1 + \log^+\frac{|f|}{\alpha}\right) + \frac{|f|}{\alpha}\log^+\alpha.$$

Since for every $t > 0$, $t \leq 2t\,(1 + \log^+ t)$, we find that

$$\frac{|f(x)|}{\alpha} \leq 2\frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right).$$

Combining (5.7) and (5.8), then replacing $|f|$ by $\frac{|f|}{\alpha}$

$$\begin{aligned}
|E_\alpha| &\leq \frac{C_1C_2}{\alpha}\int_{\mathbb{R}^3} |f(x)|(1 + \log^+ |f(x)|)\,\mathrm{d}x \\
&\leq C_1C_2\int_{\mathbb{R}^3} \frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right)\mathrm{d}x + \log^+(\alpha)\int_{\mathbb{R}^3} \frac{|f(x)|}{\alpha}\,\mathrm{d}x \\
&\leq C_1C_2(1 + 2\log^+ \alpha)\int_{\mathbb{R}^3} \frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right)\mathrm{d}x \\
&= C\int_{\mathbb{R}^3} \frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right)\mathrm{d}x\,.
\end{aligned}$$

□

The idea for the construction is based on finding a basis where Proposition 5.3.2 does not hold. Córdoba already showed that the conjecture holds for the basis $\{[s, t, \varphi(s, t)] : s > 0, t > 0\}$ so we need to make the basis more complicated if we want to find a counterexample.

Let $\mathcal{B} = \{[s,\, t\varphi(s),\, t\psi(s)] \,:\, s, t > 0\}$ be a basis where φ and ψ are two positive, continuous, and increasing functions. Note that $\mathcal{B}$ satisfies the conditions in Zygmund's Conjecture. Set $x = s$, $y = t\varphi(s)$, and $H(x, y) = y\frac{\psi(x)}{\varphi(x)}$, then

$$\mathcal{B} = \{[x,\, y,\, H(x, y)] \,:\, x, y > 0\}.$$

Since H is clearly increasing in y, Proposition 5.3.2 applies in all regions where H is monotone in x. The proposition would then apply globally if the monotonicity in x changes only finitely many times. To see this, suppose that $H(\,\cdot\,, y)$ changes from increasing to decreasing (or vice–versa) only at the x–points $0 < c_1 < \cdots < c_N < \infty$. Then the positive x–axis splits into $N + 1$ disjoint intervals $(0, c_1], (c_1, c_2], \ldots, (c_N, \infty)$, on each of which H is monotone in x. Denote the corresponding subcollections of rectangles by $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_N$. Proposition 5.3.2 then applies to each $\mathcal{B}_i$. Note that $M_{\mathcal{B}}f = \max_{0\leq i\leq N} M_{\mathcal{B}_i}f$ and thus that

$$\{M_{\mathcal{B}}f > \alpha\} = \{\max_{0\leq i\leq N} M_{\mathcal{B}_i}f > \alpha\} = \bigcup_{i=0}^{N}\{M_{\mathcal{B}_i}f > \alpha\}\,.$$

By subadditivity of the Lebesgue measure,

$$|M_{\mathcal{B}_i}f > \alpha\}| \leq \sum_{i=0}^{N}|\{M_{\mathcal{B}_i}f > \alpha\}| \leq (N + 1)\,C\int_{\mathbb{R}^d} \frac{|f(x)|}{\alpha}\left(1 + \log^+ \frac{|f(x)|}{\alpha}\right)\mathrm{d}x\,,$$

and thus Proposition 5.3.2 holds globally.

This suggests the next step in constructing a counterexample to Zygmund's Conjecture. We want to find functions φ and ψ such that H changes monotonicity infinitely many times. To this end, Soria presents the following lemma.

Lemma 5.3.3. *Let $\varphi(s) = e^{-6/s}$, $s > 0$. For $k \in \mathbb{Z}$, let $\psi(2^k) = \psi(\frac{3}{2}2^k) = \varphi(2^k)$ and define $\psi(s)$ by linear interpolation from these points. Then, given $0 \leq x \leq \alpha \leq 1$, there exists $s \in [x, 2x]$ such that $\psi(s)/\varphi(s) = \alpha$.*

Proof. Define $g(s) := \psi(s)/\varphi(s)$. Then, $g(2^k) = 1$ and $g(\frac{3}{2}2^k) = \exp\Big(-1/2^{k-1}\Big) \leq 2^{k-1}$, where we used the estimation $e^{-x} \leq \frac{1}{x}$ for all $x > 0$. Given $0 \leq x \leq \alpha \leq 1$, let j be the unique non–positive integer such that $2^{j-1} < x \leq 2^j$. Set $I_1 = [\frac{3}{2}2^{j-1}, 2^j]$ and $I_2 = [2^j, \frac{3}{2}2^j]$. We then have two possibilities: either $I_1 \subset [x, 2x]$ or $I_2 \subset [x, 2x]$. In either case, since

$$g(2^j) = 1 \geq \alpha \geq x \geq 2^{j-1} \geq g\left(\frac{3}{2}2^j\right) > g\left(\frac{3}{2}2^{j-1}\right),$$

the result follows from the Intermediate Value Theorem. □

We have now almost constructed Soria's counterexample. The only thing remaining is showing that Zygmund's Conjecture fails for the basis

$$\mathcal{B} := \{[s,\, t\varphi(s),\, t\psi(s)] \,:\, t, s > 0\}, \tag{5.9}$$

where φ and ψ are as defined in Lemma 5.3.3. To do this, we use the concept of one basis *majorising* another.

Definition 5.3.4. *Given two bases, $\mathcal{B}_1$ and $\mathcal{B}_2$, we say that $\mathcal{B}_1$* ***majorises*** *$\mathcal{B}_2$ if there exists a constant C such that for all $R \in \mathcal{B}_2$ we can find $R' \in \mathcal{B}_1$ with $R \subseteq R'$ and $|R'| \leq C|R|$.*

Remark 5.3.5. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases such that $\mathcal{B}_1$ majorises $\mathcal{B}_2$. It immediately follows from Definition 5.3.4 and the definitions of $M_{\mathcal{B}_1}f$ and $M_{\mathcal{B}_2}f$, that $M_{\mathcal{B}_2}f(x) \leq C\, M_{\mathcal{B}_1}f(x)$ for all $x \in \mathbb{R}^d$, for all $f \in L^1_{\text{loc}}$.

Take the basis $\mathcal{B}$ as in (5.9) and define

$$\mathcal{B}' := \{[x, y, z] : 0 \leq x \leq z \leq y \leq 1\} \tag{5.10}$$

We want to show that $\mathcal{B}$ majorises $\mathcal{B}'$. Let R' be an element of $\mathcal{B}'$ with dimensions $x \times y \times z$. Set $\alpha = z/y \leq 1$. Then, $x \leq \alpha \leq 1$ and from Lemma 5.3.3 we can pick $x \leq s \leq 2x$ such that $\psi(s)/\varphi(s) = z/y$. We now take $t = y/\varphi(s)$. Thus, $t\psi(s) = z$. With the choice, $s = x$, $t = y/\varphi(s)$, and $\psi(s)/\varphi(s) = z/y$, we find that the rectangle R with dimensions $s \times t\varphi(s) \times t\psi(s)$ contains R' since side s of R has length $s = x$, side $t\varphi(s)$ has length $t\varphi(s) = \frac{y}{\varphi(s)}\varphi(s) = y$, and side $t\psi(s)$ has length $t\psi(s) = \frac{y}{\varphi(s)}\psi(s) = y\frac{z}{y} = z$. In particular, $R' \subseteq R$ and $|R| = |R'|$ so the bases $\mathcal{B}$ and $\mathcal{B}'$ satisfy Definition 5.3.4 with equality and constant $C = 1$. This means that we have

$$\frac{1}{|R|}\int_R f(y)\,\mathrm{d}y = \frac{1}{|R'|}\int_{R'} f(y)\,\mathrm{d}y\,. \tag{5.11}$$

It now remains to show that we can relate $\frac{1}{|R'|}\int_{R'} f(y)\,\mathrm{d}y$ for some $R' \in \mathcal{B}'$ to the same expression for some $R'' \in \mathcal{B}''$ where we define

$$\mathcal{B}'' := \{[x, y, z] : x, y, z \in [0,1]^3\}\,. \tag{5.12}$$

There are two important observations to be made when relating these integrals. The first is that, for any $I'' \in \mathcal{B}''$, there exists a coordinate reordering, $\tau(I'') : [0,1]^3 \to [0,1]^3$, such that $I := \tau(I'') \in \mathcal{B}'$ and that if $x \in I''$, then $\tau(x) \in I$. The second is that such a re–ordering τ, is a permutation and is therefore measure–preserving under the Lebesgue measure.

There are six possible re–orderings between $\mathcal{B}''$ and $\mathcal{B}'$. We denote them as τ_i, where $1 \leq i \leq 6$. Let $R'' \in \mathcal{B}''$ and set $R' = \tau_i(R'')$. Because each τ_i is measure–preserving,

$$\frac{1}{|R''|}\int_{R''} |f(u)|\,\mathrm{d}u = \frac{1}{|R'|}\int_{R'} |f \circ \tau_i^{-1}(v)|\,\mathrm{d}v.$$

We know from sharpness of the JMZ Theorem that if we let $S = [0,1]^3$ and $\mathcal{B}''$ as defined in (5.12), then there exists a function $f \in L(\log^+ L)^1(S)$ such that the integral of f fails to be strongly differentiable almost everywhere. This means that there must also exist a function $g := f \circ \tau_i \in L(\log^+ L)^1(S)$ such that, if we let $\mathcal{B}'$ as in (5.10), the integral of g fails to be strongly differentiable almost everywhere.

In turn, this implies that, even though $\mathcal{B}$ depends on only two parameters, there exists a function in $L(\log^+ L)^1(S)$ such that its integral fails to be strongly differentiable almost everywhere, with respect to the basis $\mathcal{B}$. This concludes the counterexample to Conjecture 5.2.1.

The counterexample constructed for $\mathbb{R}^3$ can now easily be generalised to $\mathbb{R}^d$ for $d \geq 4$. There are multiple ways to do this, but the most obvious one would be to take φ and ψ as in Lemma 5.3.3 and define the basis

$$\mathcal{B} := \{[s, t\varphi(s), t\psi(s), t_4, t_5, \ldots, t_d] \,:\, s,\, t,\, t_i > 0\}.$$

From the arguments used earlier, we know that $\mathcal{B}$ behaves as the basis of all rectangles in $\mathbb{R}^d$, even though it depends only on $d-1$ parameters. This contradicts Zygmund's Conjecture for $d \geq 4$. Together these constructions conclude the counterexamples to Zygmund's Conjecture for $d \geq 3$ and show that, in its most general form, the conjecture does not hold.

5.4. Counterexamples by G. Rey

Where some counterexamples were already found by F. Soria in 1984, Guillermo Rey found additional counterexamples much more recently. In 2020, Rey published an article in which he builds on Soria's work to construct more counterexamples to Zygmund's conjecture [14]. Specifically, he shows that the bases of rectangles

$$\mathcal{B} = \{[2^{\varphi_1(s,t)}, \ldots, 2^{\varphi_d(s,t)}] : s, t > 0\}$$

for all dimensions $d \geq 3$ do not differentiate $L(\log^+ L)^{d-2} \subseteq L(\log^+ L)^1$. This means that Zygmund's Conjecture (Conj. 5.2.1) does not hold for $k = 2$ in all dimensions $d \geq 3$.

The construction of Rey's counterexamples is based on dyadic intervals (see Section 3.1) so the statement of the conjecture needs to be slightly modified.

Conjecture 5.4.1 (Zygmund's Conjecture, Dyadic). *Let $f \in L(\log^+ L)^{k-1}(S)$ with $S \subseteq \mathbb{R}^d$ and $1 \leq k \leq d$. Define the basis $\mathcal{B} := \{[2^{\varphi_1(n)}, \ldots, 2^{\varphi_d(n)}] : n \in \mathbb{Z}^k\}$, where each $\varphi_i : \mathbb{Z}^d \to \mathbb{Z}$ is non–decreasing in each variable. Then, for $I \in \mathcal{B}$,*

$$\lim_{\substack{\delta(I)\to 0 \\ I \ni x}} \frac{1}{|I|} \int_I |f(y)| \, \mathrm{d}y = f(x) \quad \text{for almost all} \quad x \in S\,.$$

To see that this statement is equivalent to Conjecture 5.2.1 we need the One–Third Trick (Lem. 3.1.4). If we let the shifted dyadic grid $\mathcal{D}^{(\boldsymbol{\sigma})}$ be as in Lemma 3.1.4, then by similar arguments the One–Third Trick can be applied to axis–aligned rectangles I.

First note that any rectangle $I_d \in \mathcal{B}_d := \{[2^{\varphi_1(n)}, \ldots, 2^{\varphi_d(n)}] : n \in \mathbb{Z}^k\}$ in the dyadic version of Zygmund's Conjecture is automatically also a rectangle in some $\mathcal{B}$ as defined in Conjecture 5.4.1. Hence, $\mathcal{B}_d \subseteq \mathcal{B}$ and thus if Conjecture 5.2.1 is true, then so is Conjecture 5.4.1.

Second, every shifted dyadic grid $\mathcal{D}^{(\boldsymbol{\sigma})}$ is contained in the basis of dyadic rectangles $\mathcal{B}_d$. The only step that remains is applying the One–Third Trick. For every $I \in \mathcal{B}$, there exists an $R \in \mathcal{D}^{(\boldsymbol{\sigma})} \subseteq \mathcal{B}_d$ such that $I \subseteq R$ and $|R| \leq C_d\,|I|$, i.e. I is contained in a dyadic set of comparable size. Assume that Conjecture 5.4.1 holds. By the argument we just gave,

$$\left| \int_I f(y)\,\mathrm{d}y - f(x) \right| \leq \left| \int_I f(y)\,\mathrm{d}y - \int_R f(y)\,\mathrm{d}y \right| + \left| \int_J f(y)\,\mathrm{d}y - f(x) \right|.$$

The second term goes to zero as $\delta(J) \to 0$ by assumption. The first term is controlled since $I \subseteq R$ and their sizes differ only by a finite constant factor; as both diameters go to zero, their averages converge together. Thus, if Conjecture 5.4.1 holds, then so does Conjecture 5.2.1. Putting the two implications together yields the equivalence of the two statements of Zygmund's Conjecture.

The following simple lemma is needed for the construction of Rey's counterexamples.

Lemma 5.4.2. *For any function $\Phi : \mathbb{Z} \to \mathbb{Z}$, there exists a function $\varphi : \mathbb{Z}^2 \to \mathbb{Z}$ that satisfies*

i. $\varphi(n, -n) = \Phi(n)$ *for all* $n \in \mathbb{Z}$,

ii. φ *is non–decreasing in each variable.*

Proof. We split the grid of integers, $\mathbb{Z}^2$, into two regions. We consider all points on the right–hand side of the line $(n, -n)$, also called the *anti–diagonal*, as one region, and all points on the left–hand side of this line as the other region.

Take $n_1, n_2 \in \mathbb{Z}$ where $n_1 \geq -n_2$, then (n_1, n_2) is on the right–hand side of the anti–diagonal, and set

$$\varphi(n_1, n_2) = \max_{n \in [-n_2, n_1]} \Phi(n) .$$

As n_1 or n_2 increases, so does the set over which the maximum is taken, and thus φ, on the right–hand side of the anti–diagonal, is a non–decreasing function in each variable.

On the left–hand side of the anti–diagonal, so for $n_1 \leq -n_2$, we set

$$\varphi(n_1, n_2) = \min_{n \in [n_1, -n_2]} \Phi(n) ,$$

and by a similar argument, φ is also a non–decreasing function in both variables on the left–hand side of the anti–diagonal. Thus, the function $\varphi : \mathbb{Z}^2 \to \mathbb{Z}$, as we have constructed it, exactly satisfies the conditions in the lemma. □

With Lemma 5.4.2 we can prove that for $k = 2$ there exist functions $\varphi_i : \mathbb{Z}^k \to \mathbb{Z}$ for $1 \leq i \leq d$ that are non–decreasing in each variable such that, if we take $I \in \mathcal{B}$, then there exists a function $f \in L(\log^+ L)^{d-2}$ for which the integral fails to be strongly differentiable almost everywhere.

Define the basis

$$\mathcal{B}(2) := \{[2^{\varphi_1(n_1,n_2)}, \ldots, 2^{\varphi_d(n_1,n_2)}] \,:\, n_1, n_2 \in \mathbb{Z}\} .$$

Choose any bijection $\psi : \mathbb{Z} \to \mathbb{Z}^d$ and set for $1 \leq i \leq d$

$$\Phi_i(n) = \psi_i(n) \quad \text{for all} \;\; n \in \mathbb{Z},$$

where $\Phi_i : \mathbb{Z} \to \mathbb{Z}$. Now apply Lemma 5.4.2 to Φ_i and denote the resulting function by φ_i, then $\varphi_i : \mathbb{Z}^2 \to \mathbb{Z}$ and $\varphi_i(n, -n) = \Phi_i(n) = \psi_i(n)$ for every $n \in \mathbb{Z}$. By construction we have, for every integer $n \in \mathbb{Z}$,

$$(\varphi_1(n, -n), \ldots, \varphi_d(-n, n)) = (\Phi_1(n), \ldots, \Phi_d(n)) = \psi(n) .$$

Because ψ is a bijection from $\mathbb{Z}$ to $\mathbb{Z}^d$, the set of all d-tuples $(\varphi_1(n, -n), \ldots, \varphi_d(-n, n))$ as n ranges over $\mathbb{Z}$ is exactly the whole grid $\mathbb{Z}^d$:

$$\{[\varphi_1(n, -n), \ldots, \varphi_d(n, -n)] \,:\, n \in \mathbb{Z}\} = \mathbb{Z}^d .$$

Take an arbitrary dyadic rectangle $[2^{m_1}, \ldots, 2^{m_d}]$ with $(m_1, \ldots, m_d) \in \mathbb{Z}^d$. Because of the above equality, there exists an integer n such that

$$(\varphi_1(n, -n), \ldots, \varphi_d(-n, n)) = (m_1, \ldots, m_d) .$$

Setting $n_1 = n$ and $n_2 = -n$ we obtain

$$[2^{\varphi_1(n_1,n_2)}, \ldots, 2^{\varphi_d(n_1,n_2)}] = [2^{m_1}, \ldots, 2^{m_d}]$$

and thus the basis $\mathcal{B}(2)$ is the collection of all dyadic rectangles in $\mathbb{R}^d$. It follows directly from sharpness of the Jessen–Marcinkiewicz–Zygmund Theorem that there exists a function $f \in L(\log^+ L)^{d-2}$, for which the integral of f fails to be strongly differentiable almost everywhere. And thus, the constructed basis is a counterexample to Zygmund's Conjecture for $k = 2$.

The result for $k = 2$ can be generalised to $2 \leq k < d$. Define the basis

$$\mathcal{B}(k) := \{[2^{\vartheta_1(n)}, \dots, 2^{\vartheta_d(n)}] : n \in \mathbb{Z}^k\}, \tag{5.13}$$

where each ϑ_i is dependent on k variables. For $k > 2$ we can just define

$$\vartheta_i(n_1, \dots, n_k) := \varphi_i(n_1, n_2), \qquad n_1, \dots, n_k \in \mathbb{Z},$$

where φ_i is as we constructed earlier. This yields the same result as before; the basis $\mathcal{B}(k)$ as defined in (5.13) is the collection of all dyadic rectangles in $\mathbb{R}^d$. It again follows directly from sharpness of the Jessen–Marcinkiewicz–Zygmund Theorem, that there exists a function $f \in L(\log^+ L)^{d-2}$, for which the integral of f fails to be strongly differentiable almost everywhere if we take $I \in \mathcal{B}(k)$.

To see how these constructions form counterexamples to Zygmund's Conjecture, we reason as follows. For $2 \leq k < d$, note that $k - 1 \leq d - 2$ and thus that $L(\log^+ L)^{d-2} \subseteq L(\log^+ L)^{k-1}$. This means that any $f \in L(\log^+ L)^{d-2}$, f must also be in $L(\log^+ L)^{k-1}$. We know that, if we let $I \in \mathcal{B}(k)$ as defined in (5.13), then there exists a function $f \in L(\log^+ L)^{d-2}$ such that its integral fails to be strongly differentiable almost everywhere. Taking this same function f, it follows that if we let $I \in \mathcal{B}(k)$, then there exists a function $f \in L(\log^+ L)^{k-1}$ such that its integral fails to be strongly differentiable almost everywhere. This proves that the conjecture, in its most general form, fails for all $d \geq 3$ and $2 \leq k < d$.

5.5. Examples by Córdoba and Soria

Although F. Soria and G. Rey showed in [12, 14] that Zygmund's Conjecture does not hold in its most general form, there might be conditions that we can place on the non–decreasing functions φ_i in the basis $\mathcal{B}$, such that the conjecture does hold.

As we stated at the start of this chapter, there are bases for which the conjecture does in fact hold. The first examples were constructed by A. Córdoba [13]. He showed that the conjecture holds in $\mathbb{R}^3$ for $\mathcal{B} := \{[s,\, t,\, \varphi(s,t)] \,:\, s,\, t > 0\}$, where $\varphi(s,t)$ is any positive function that is non–decreasing in each variable separately.

Soria extends this result in [12] to some bases of the form

$$\mathcal{B} := \{[\varphi_1(s,t),\, \varphi_2(s,t),\, \varphi_3(s,t)] \,:\, s,\, t > 0\}\,. \tag{5.14}$$

For the conjecture to hold for such bases, it is necessary to impose some relationships between the different φ_i. Soria does this as follows.

Let $\varphi_1(s,t)$, $\varphi_2(s,t)$, $\varphi_3(s,t)$ be three smooth functions that increase in each variable separately. Define $F(s,t) := (\varphi_1(s,t),\, \varphi_2(s,t))$ and assume that F is one–to–one in some region $A \subseteq (0,\infty) \times (0,\infty)$. For $(x,y) \in F(A)$ we can define $H(x,y) = \varphi_3(F^{-1}(x,y))$ so that $\varphi_3(s,t) = H(\varphi_1(s,t),\, \varphi_2(s,t))$. By the chain rule, and denoting the gradient by ∇, we obtain

$$\nabla\varphi_3 = \nabla H \cdot \begin{pmatrix} \dfrac{\partial\varphi_1}{\partial s} & \dfrac{\partial\varphi_1}{\partial t} \\ \dfrac{\partial\varphi_2}{\partial s} & \dfrac{\partial\varphi_2}{\partial t} \end{pmatrix}\,.$$

We assumed that each φ_i increased on each variable, so the partial derivatives are all positive. Therefore, at each point, at least one partial derivative of H must be non–negative. If we choose the set $A \subseteq (0,\infty) \times (0,\infty)$ so that the partial derivatives of H do not change their signs on A, then we can apply Proposition 5.3.2 to the basis $\mathcal{B} := \{[\varphi_1(s,t),\, \varphi_2(s,t),\, \varphi_3(s,t)] \,:\, (s,\, t) \in A\}$.

Constructing an explicit example from this proposition can be very straightforward. We could, for instance, let φ_i be monomials, i.e. let $\varphi_i(s,t) := s^{\alpha_i}\, t^{\beta_i}$ for some constants $\alpha_i,\, \beta_i \geq 0$. To see this, we compute the Jacobian:

$$J(\varphi_i, \varphi_j) = \det \begin{pmatrix} \dfrac{\partial \varphi_i}{\partial s} & \dfrac{\partial \varphi_i}{\partial t} \\ \dfrac{\partial \varphi_j}{\partial s} & \dfrac{\partial \varphi_j}{\partial t} \end{pmatrix} = \det \begin{pmatrix} \alpha_i\, s^{\alpha_i - 1} t^{\beta_i} & \beta_i\, s^{\alpha_i} t^{\beta_i - 1} \\ \alpha_j\, s^{\alpha_j - 1} t^{\beta_j} & \beta_j\, s^{\alpha_j} t^{\beta_j - 1} \end{pmatrix} = (\alpha_i \beta_j - \beta_i \alpha_j)\, s^{\alpha_i + \alpha_j - 1}\, t^{\beta_i + \beta_j - 1}\,.$$

Thus the Jacobian $J(\varphi_i, \varphi_j)$ only vanishes in the first quadrant when $\alpha_i\, \beta_j - \alpha_j\, \beta_i = 0$. For the i and j where the Jacobian does vanish, there exists a constant $m \geq 0$ such that $\varphi_i(s,t) = [\varphi_j(s,t)]^m$. This means that the i–side of every rectangle in $\mathcal{B}$ is a fixed monotone function of the j–side. If we have a basis for which that is the case, we reason as follows.

Suppose that there are $1 \leq i < j \leq 3$ such that $\alpha_i\, \beta_j - \alpha_j\, \beta_i = 0$. Let $m \geq 0$ be such that $\varphi_i(s,t) = [\varphi_j(s,t)]^m$. Let $R = R(s,t)$ and $R' = R(s',t')$ both be rectangles in $\mathcal{B}$ as defined in (5.14). Assume that we have $\varphi_i(s,t) > \varphi_i(s',t')$ and note that the map $x \mapsto x^{1/m}$ is non–decreasing for $m \geq 0$. Since $\varphi_j(s,t) = [\varphi_i(s,t)]^{1/m}$ for some $m \geq 0$ and arbitrary $s, t \in A$, we have $\varphi_j(s,t) > \varphi_j(s',t')$. In words; whenever R is longer than R' in the i–direction, it is also longer in the j–direction. Hence, the ordering condition in Proposition 5.3.2 is satisfied and Zygmund's Conjecture holds for the basis $\mathcal{B}$ with the functions $\varphi_i(s,t) := s^{\alpha_i}\, t^{\beta_i}$.

This chapter summarises results regarding Zygmund's Conjecture. While we have to conclude that the conjecture does not hold when it is stated in its most general form, there is still a good chance that a slightly more specific statement holds. There are simply too many bases for which we do not know whether or not they satisfy the conjecture. The example given by Soria in [12] indicates that with some conditions on the functions φ_i that determine the sides of the rectangles in $\mathcal{B}$, the conjecture may very well hold.

So even though we cannot end this chapter by providing an exhaustive list of bases that do and do not differentiate the space $L(\log^+ L)^k$, we can end on the positive note that it is likely that the current list of bases will continue to grow and that we might even end up with a version of the conjecture that does hold.

Chapter 6

Concluding Remarks and Outlook

The aim of this thesis was to provide a complete and consistent text on the Jessen–Marcinkiewicz–Zygmund (JMZ) Theorem, and the later extension due to Zygmund, at a level accessible to advanced undergraduates. We began by recalling the essential measure–theoretic background, then developed the Hardy–Littlewood maximal operator and the Lebesgue Differentiation Theorem (LDT).

After this, we covered the JMZ Theorem and the results leading up to it. Where most details were omitted in the original 1935 article [4], we gave complete proofs to all intermediate results. In addition, we provided a proof of the sharpness of the JMZ Theorem for all $d \geq 2$.

We continued by covering Zygmund's 1967 extension [5] of the JMZ Theorem. Here we again filled gaps whose details are often omitted in research papers, to ensure that an undergraduate reader can follow the logic without having to rely on external sources. Moving on to Zygmund's Conjecture [5], we reconstructed the counterexamples by F. Soria [12] and G. Rey [14], providing the details of the reasoning behind each step. We ended the discussion of Zygmund's work by presenting positive results on the conjecture.

Overall, this thesis offers a thorough presentation of the Jessen–Marcinkiewicz–Zygmund Theorem and the extension found by Zygmund. These deep results and their details are often scattered across specialist papers, which use varying notation. By providing the necessary preliminaries, complete proofs, and unified notation, we make the material fully accessible to undergraduate students and other interested mathematicians. We expect this to offer a way to bridge the gap between a course in real analysis and the available research regarding differentiation theorems and maximal operators.

The discussion of Zygmund's Conjecture showed us that, even though the conjecture is known to fail when stated in its most general form, there is hope that a more precise re–statement of the conjecture might hold. The area remains open to research and while it was first conjectured by Zygmund in 1967, the 2020 article by Rey [14] indicates that many questions are still unsolved. Finding more examples and counterexamples to the Zygmund Conjecture is an interesting topic for future research as it helps us better understand the precise geometry behind maximal operators and their applications.

Bibliography

[1] Loukas Grafakos. *Fundamentals of Fourier Analysis.* Springer Cham, 2024.

[2] Emiel Lorist and Mark Veraar. *Harmonic Analysis - Lecture Notes.* TU Delft, 2025.

[3] G. B. Folland. *Real analysis : modern techniques and their applications (2nd ed).* John Wiley & Sons, 1999.

[4] Børge Jessen, Józef Marcinkiewicz, and Antoni Zygmund. "Note on the differentiability of multiple integrals". ENG. In: *Fundamenta Mathematicae* 25.1 (1935), pp. 217–234.

[5] Antoni Zygmund. "A note on the differentiability of integrals". eng. In: *Colloquium Mathematicum* 16.1 (1967), pp. 199–204. URL: http://eudml.org/doc/262993.

[6] N. L. Carothers. *Real Analysis.* Cambridge University Press, 2000.

[7] Mark Veraar and Emiel Lorist. *Lecture Notes AM2090, Real Analysis.* TU Delft, 2023.

[8] Terence Tao. *Lecture notes 3 for 247A.* Department of Mathematics, UCLA. URL: https://www.math.ucla.edu/~tao/247a.1.06f/notes3.pdf.

[9] Richard Lechner. "The One-Third-Trick and Shift Operators". In: *Bulletin of the Polish Academy of Sciences Mathematics* 61.3 (2013), pp. 219–238. ISSN: 1732-8985. DOI: 10.4064/ba61-3-4. URL: http://dx.doi.org/10.4064/ba61-3-4.

[10] G. H. Hardy and J. E. Littlewood. "A maximal theorem with function-theoretic applications". In: *Acta Mathematica* 54 (1 1930), pp. 81–116.

[11] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, 1923.

[12] Fernando Soria. "Examples and Counterexamples to a Conjecture in the Theory of Differentiation of Integrals". In: *Annals of Mathematics* 123.1 (1986), pp. 1–9. ISSN: 0003486X, 19398980. URL: http://www.jstor.org/stable/1971350 (visited on 05/08/2025).

[13] Antonio Córdoba. "The Theorem $s \times t \times \varphi(s, t)$". In: *Mittag-Leffler Institut* 9 (1978).

[14] Guillermo Rey. *Another counterexample to Zygmund's conjecture.* 2020. arXiv: 2003.11392 [math.CA]. URL: https://arxiv.org/abs/2003.11392.

Appendix A

Omitted Proofs

A.1. Explicit Proofs for Zygmund 1967

Lemma 5.1.2

Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^k(S)$ and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths. Then $M_S f \in L^1(S)$ and

$$\int_S M_S f(y)\,\mathrm{d}y \leq A \int_S |f(y)|(1 + \log^+|f(y)|)^k\,\mathrm{d}y + B,$$

where A and B depend only on d.

Proof. We may assume that $f \geq 0$ since $M_S f(x) = M_S|f|(x)$ and split f into dyadic intervals. For each $m = 1, 2, \ldots$, we define

$$E_m := \{2^{m-1} \leq f \leq 2^m\}, \quad f_m = f \cdot \mathbb{1}_{E_m}.$$

For $m = 0$ we define $E_0 := \{f < 1\}$ and $f_0 = f \cdot \mathbb{1}_{E_0}$. Then $f = \sum_{m=0}^{\infty} f_m$ and by subadditivity, $M_S f \leq \sum_{m=0}^{\infty} M_S f_m$.

By Lemma 5.1.1,

$$\int_S M_S f_m(y)\,\mathrm{d}y \leq \left(\int_S (M_S f_m(y))^p\,\mathrm{d}y\right)^{1/p} \leq \frac{A}{(p-1)^k}\left(\int_S (f_m(y))^p\,\mathrm{d}y\right)^{1/p} \leq \frac{A}{(p-1)^k} 2^m |E_m|^{1/p}.$$

Set $p = 1 + \frac{1}{m+1}$ for $m = 0, 1, \ldots$, then $\frac{1}{p-1} = m + 1$ and thus

$$\int_S M_S f_m(y)\,\mathrm{d}y \leq 3(m+2)^k 2^m |E_m|^{(m+1)/(m+2)}.$$

Summing over m yields

$$\int_S M_S f(y)\,\mathrm{d}y \leq C \sum_{m=0}^{\infty} (m+1)^k 2^m |E_m|^{(m+1)/(m+2)} \;. \tag{A.1}$$

Since $2^{m-1} < f \leq 2^m$ on E_m, we have $f(y)(1 + \log^+ f(y))^k \geq 2^{m-1}(m-1)^k (\log 2)^k$ on E_m. Thus

$$\int_S f(y)(1 + \log^+ f(y))^k\,\mathrm{d}y = \sum_{m=0}^{\infty}\left(\int_{E_m} f(y)(1 + \log^+ f(y))^k\,\mathrm{d}y\right) \geq \frac{1}{2}(\log 2)^k \sum_{m=0}^{\infty} (m-1)^k 2^m |E_m|. \tag{A.2}$$

In order to be able to compare (A.1) and (A.2), we need to relate $|E_m|^{(m+1)/(m+2)}$ to $|E_m|$. We split the indices into two classes, one for which the measure of the set is large, $|E_m| \geq 3^{-m}$, and one for which the measure is small, $|E_m| < 3^{-m}$. For m where $|E_m| \geq 3^{-m}$ we have

$$|E_m|^{(m+1)/(m+2)} = |E_m||E_m|^{-1/(m+2)} \leq 3\,|E_m|.$$

If $|E_m| < 3^{-m}$, then

$$(m+1)^k 2^m |E_m|^{(m+1)/(m+2)} \leq (m+1)^k 2^m 3^{-m} \leq \tilde{C}(2/3)^m.$$

The sum of $(m+1)^k 2^m |E_m|^{(m+1)/(m+2)}$ over all m for which $|E_m| < 3^{-m}$ is finite and can be estimated by some constant.

Combining the two classes we obtain, with finite constants $\tilde{A}$ and $\tilde{B}$,

$$C \sum_{m=0}^{\infty} (m+1)^k 2^m |E_m|^{(m+1)/(m+2)} \leq \tilde{A}\Big(\sum_{m=0}^{\infty} (m-1)^k 2^m |E_m| \Big) + \tilde{B}\,.$$

We can now relate (A.1) and (A.2) to find

$$\begin{aligned}\int_S M_S f(y)\,\mathrm{d}y &\leq C \sum_{m=0}^{\infty} (m+1)^k 2^m |E_m|^{(m+1)/(m+2)} \\ &\leq \tilde{A}\Big(\sum_{m=0}^{\infty} (m-1)^k 2^m |E_m| \Big) + \tilde{B} \\ &\leq A \int_S f(y)(1+\log^+ f(y))^k \,\mathrm{d}y + B\,.\end{aligned}$$

□

Theorem 5.1.4

Let $f : S \to \mathbb{C}$ be a measurable function and $S \subseteq \mathbb{R}^d$. If $f \in L(\log^+ L)^{k-1}(S)$ and $I \subseteq S$ has $1 \leq k \leq d$ distinct side lengths, then the integral of $f(x)$ is strongly differentiable with respect to I at almost every $x \in S$, and the derivative is equal to $f(x)$.

Proof. Apply (4.3) to the function λf where $f \in L(\log^+ L)^{k-1}(S)$, and $\lambda > 0$ is a constant, then by Lemma 5.1.3

$$\int_S m_S f(y)\,\mathrm{d}y \leq A \int_S |f(y)|(1+\log^+|\lambda f(y)|)^{k-1}\,\mathrm{d}y + \frac{B}{\lambda}.$$

Let $\varepsilon > 0$ be arbitrary and choose λ so large that $B/\lambda < \frac{1}{2}\varepsilon$. By Lusin's Theorem (Thm. 2.1.10), f can be decomposed as $f(y) = \varphi(y) + \psi(y)$ where φ is continuous and ψ satisfies

$$\int_S |\psi(y)|\,\mathrm{d}y < \varepsilon, \tag{A.3}$$

$$A\int_S |\psi(y)|(1+\log^+|\lambda\psi(y)|)^{k-1}\,\mathrm{d}y + \frac{B}{\lambda} < \varepsilon. \tag{A.4}$$

We apply Lemma 5.1.3 to $\lambda\psi$ and divide by λ,

$$\int_S m_S \psi(y)\,\mathrm{d}y \leq A\int_S |\psi(y)|(1+\log^+|\lambda\psi(y)|)^{k-1}\,\mathrm{d}y + \frac{B}{\lambda} < \varepsilon. \tag{A.5}$$

For the last steps of the proof, we need to define two sets, $E_1(\varepsilon) := \{y \in S : |\psi(y)| > \sqrt{\varepsilon}\}$ and $E_2(\varepsilon) := \{y \in S : m_S\psi(y) > \sqrt{\varepsilon}\}$, and denote their union by $E(\varepsilon)$. We apply Chebyshev (Prop. 2.2.9) with $p = 1$ to these sets. Recall that $||f||_{p,\infty} = \sup_{t>0}\ t|\{|f| > t\}|^{1/p}$. Then,

$$|E_1(\varepsilon)| = |\{|\psi(y)| > \sqrt{\varepsilon}\}| \leq \frac{1}{\sqrt{\varepsilon}} \int_S |\psi(y)|\,\mathrm{d}y < \frac{\varepsilon}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}.$$

Similarly for $E_2(\varepsilon)$,

$$|E_2(\varepsilon)| = |\{m_S\psi(y) > \sqrt{\varepsilon}\}| \leq \frac{1}{\sqrt{\varepsilon}} \int_S m_S\psi(y)\,\mathrm{d}y < \frac{\varepsilon}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}.$$

Thus, $|E(\varepsilon)| \leq |E_1(\varepsilon)| + |E_2(\varepsilon)| < 2\sqrt{\varepsilon}$. With the decomposition of f, and $I \ni x$ any axis-aligned rectangle with at most k different side-lengths, we write,

$$\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x) = \left[\frac{1}{|I|}\int_I \varphi(y)\,\mathrm{d}y - \varphi(x)\right] + \left[\frac{1}{|I|}\int_I \psi(y)\,\mathrm{d}y - \psi(x)\right].$$

For $n \in \mathbb{N}$ set $\varepsilon_n := 4^{-n}$ and $E_n := E(\varepsilon_n)$. Then

$$|E_n| \leq 2\sqrt{\varepsilon_n} = 2 \cdot 2^{-n},$$

and thus $\sum_{n=1}^{\infty} |E_n| < \infty$. Define

$$S_0 := S \setminus \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} E_n = S \setminus \limsup_{n\to\infty} E_n\,.$$

The Borel-Cantelli Lemma tells us that

$$|\limsup_{n\to\infty} E_n| = 0\,,$$

so S_0 has full measure.

By definition of $\limsup E_n$, if there exists an $x \notin \limsup E_n$, then we can choose an $N = N(x)$ such that $x \notin E_n$ for all $n \geq N$. Let N be such an index. Because $x \notin E_n$ for all $n \geq N$,

$$\limsup_{\delta(I)\to 0} \left|\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x)\right| \leq m_S\psi(x) + |\psi(x)| \leq 2\sqrt{\varepsilon_n} = 2 \cdot 2^{-n}\,. \tag{A.6}$$

Recall that φ is continuous and therefore $\left(\frac{1}{|I|}\int_I \varphi(P)\,\mathrm{d}P - \varphi(P_0)\right) \to 0$ as I shrinks to x. Since the right-hand side of (A.6) tends to zero as $n \to \infty$, we find that

$$\limsup_{\delta(I)\to 0} \left|\frac{1}{|I|}\int_I f(y)\,\mathrm{d}y - f(x)\right| = 0 \qquad \text{for every} \quad x \in S_0\,. \qquad \square$$