Carleson’s Theorem on a.e. Convergence of Fourier Series

Jesús Ocáriz Gallego

Advisor: Dr. Matías Raja Baño

Faculty of Mathematics
University of Murcia
July 2015
Contents

Introduction ................................................................................. 1

1 Problem Reduction ......................................................... 13
  1.1Multiplier transference ....................................................... 13
  1.2Isometries ........................................................................ 16
  1.3Dyadic decomposition .......................................................... 20
  1.4Discretization .................................................................... 23
  1.5Last reduction ..................................................................... 29

2 The Proof of Carleson’s Theorem ........................................... 31
  2.1Idea of the proof ................................................................. 31
  2.2Technical Lemmas ............................................................... 33
  2.3The proof of Carleson’s Theorem ......................................... 34
  2.4Proof of the Density lemma .................................................. 37
  2.5Proof of the Size Lemma ....................................................... 40
  2.6Proof of the Tree lemma ....................................................... 48
Fourier’s Theorem is not only one of the most beautiful results of modern analysis, but it is said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics.

Lord Kelvin (1824-1907)

According to Dr. Antonio Córdoba in [Cór00], it is not easy to point out the breakthroughs in Analysis in the 20th century. In that article, he highlights two breakthroughs which he considers important in Analysis:

1. The Carleson Theorem about convergence of trigonometric series.
2. The Di Giorgi-Nash Theorem about the regularity of weak solutions of elliptic partial differential equations.

Among these pages of the final master project, we are going to try to prove in an understanding way the first milestone mentioned above, that is, the Carleson’s Theorem.

Let us first enunciate Carleson’s Theorem:

Carleson’s Theorem
For $f \in L^2[-\pi, \pi]$, its trigonometric series converges almost everywhere.

This statement will be understood more precisely later during the historical introduction, where the concepts of trigonometric series and convergence will be defined.

We remark that we are focused in proving Carleson’s Theorem. Consequently, no
elementary, well known or contextual result will be proved. Instead, we are going to give some reference books where the reader can find the enunciate and proof of these known results or, in some cases, we are going to cite the original articles of the contextual results. However, other results and concepts that, in my mere opinion, I consider worth knowing to understand better the proof, will be written in the final appendix.

This is the sketch we are going to follow in this introduction:

First, I would like to explain the importance and the difficulty of this problem with a historical introduction using some ideas of the article [Cór00].

Second, we are going to talk briefly about some kind of generalization of Carleson’s Theorem. Nowadays, there are still open problems related to it.

Thirdly, I am going to write the sketch of the proof we are showing in the following chapters.

Finally, I will try to answer the question if there is any simpler proof if we change the space of functions where the convergence of the trigonometric series holds. We remind that for $2\pi$ periodic functions, the following inclusions of functions spaces holds\(^1\):

\[
C^1(\mathbb{T}) \subset C(\mathbb{T}) \subset L^\infty[-\pi, \pi] \subset L^q[-\pi, \pi] \subset L^p[-\pi, \pi] \subset L^1[-\pi, \pi]
\]

where $1 < p < q < \infty$ and $C^1(\mathbb{T})$ denotes the space of $2\pi$ periodic derivable functions with continuous derivative; $C(\mathbb{T})$ denotes the space of $2\pi$ periodic continuous functions; and $L^r[-\pi, \pi]$ denotes the usual Lebesgue functions space with $1 \leq r \leq \infty$. In fact, the natural question we will ask is that if there is a simplification for functions in $C[-\pi, \pi]$.

---

\(^1\)The inclusions between the Lebesgue spaces are a clearly consequence of Hölder’s inequality and the finite Lebesgue measure of the domain $[-\pi, \pi]$. 
Historical Introduction

A trigonometric series of a function is a series of the form

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

where the coefficients depend on the function.

The trigonometric series theory started in the 18th century in relation with the vibrant string problem. Fourier was the first to assert that every periodic function \(f\) can be 'represented' in this way. For more about the relationship between Fourier series and the PDEs, the reader can follow the following reference [Cañ02].

In order to formalize Fourier's claim, we are going to say exactly the function space where we take the functions \(f\) and the meaning of 'represented'. Our functions \(f\) will belong to \(L^1[-\pi, \pi]\) (or even a better function space). Intuitively, we would like that:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

This "equality" would imply\(^2\) that for every \(f \in L^1[-\pi, \pi]\), its trigonometric series must be defined by the following coefficients:

$$a_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad n \geq 0$$
$$b_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \quad n \geq 1$$

These coefficients are known as Fourier coefficients and are well-defined because it is clearly that both \(|a_n|, |b_n| \leq \|f\|_1\).

An important fact is to describe how we sum this infinite sum. There are different concepts of this summation (classical, Abel\(^3\), Poisson), but we are interested in the more intuitive and classical one: the series \(\sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))\) is defined, if it exists, as the limit of the sequence \(S_k f(x) = \sum_{n=1}^{k} (a_n \cos(nx) + b_n \sin(nx))\), that is, \(S_k f\) are the finite sums of the trigonometric series obtained by \(f\).

\(^2\)It’s a direct consequence of the computation of integrals of type \(\int_{-\pi}^{\pi} \cos(mx) \cos(nx) dx, \int_{-\pi}^{\pi} \sin(mx) \cos(nx) dx\), and \(\int_{-\pi}^{\pi} \sin(mx) \sin(nx) dx\) if we could interchange the sum with the integration. In other words, in this way we check that these functions seen in \(L^2[-\pi, \pi]\) form an orthogonal system.

\(^3\)For example this concept of summation is important to prove the Fejer’s theorem and gets as a corollary that the trigonometric polynomials are dense in \(L^p[-\pi, \pi]\) for \(1 \leq p < \infty\). See [DC01, p. 10].
To be more precise, in what sense can we say that \( f \) coincides with its trigonometric series? In order to answer this, two problems appeared in a natural way:

**Q1.** Does the trigonometric series converge in norm to the function?

**Q2.** Does the trigonometric series converge pointwise to the function?

Let us remember how this kind of convergences are related in the function space \( L^p[-\pi, \pi] \) with \( 1 \leq p < \infty \):

- If a sequence of functions \( \{f_n\} \) converges in norm to a function \( f \), then there is a subsequence that converges pointwise almost everywhere to the function \( f \).

- If a sequence of functions \( \{f_n\} \) converges pointwise almost everywhere to a function \( f \), it is not necessary that \( \{f_n\} \) converges in norm to the function \( f \).

- If a sequence of functions \( \{f_n\} \) converges pointwise almost everywhere to a function \( f \) and is also dominated by some function \( g \in L^p[-\pi, \pi] \), then \( \{f_n\} \) converges in norm to the function \( f \). This is the Dominated Convergence Theorem.

On the one hand, the answer to **Q1** was given at the beginning of the 20th century. In 1907, F. Riesz and E. Fischer answered positively to **Q1** if \( f \in L^2[-\pi, \pi] \). Later, it was proved that it was true for \( f \in L^p[-\pi, \pi] \) with \( p > 1 \) and false with \( f \in L^1[-\pi, \pi] \).

On the other hand, it was with Carleson’s Theorem [Car66] when we got the “full” answer of **Q2**. In concrete, in the question **Q2** we can only hope to obtain pointwise convergence of the trigonometric series.

---

\( ^4 \)This proposition is related to the Riesz-Fischer theorem and its proof can be seen in [Cañ02, p. 74].

\( ^5 \)Define \( f_n(x) = n \) for \( x \in (0, 1/n] \) and \( f_n(x) = 0 \) otherwise. Then, \( \int_0^1 \lim_{n \to \infty} f_n(x) \, dx = 0 \neq 1 = \lim_{n \to \infty} \int_0^1 f_n(x) \, dx \).

\( ^6 \)In its most common form, the Riesz-Fischer Theorem says that for \( 1 \leq p < \infty \), the space \( L^p(X, \mu) \) is (sequentially) complete for all measure spaces \( (X, \Sigma, \mu) \), i.e., that every Cauchy sequence of functions in \( L^p(X, \mu) \) converges to an \( L^p(X, \mu) \)-function \( f \). This statement is far more general than the original results published by Riesz and Fischer, however, as Riesz’s result related convergence of square-summable sequences of real numbers to orthonormal systems in \( L^2[-\pi, \pi] \) while Fischer’s result proved (using more antiquated terminology and notation) the \( L^2[-\pi, \pi] \)-convergence to a function \( f \) in \( L^2[-\pi, \pi] \) for any Cauchy sequence in \( L^2[-\pi, \pi] \).

\( ^7 \)The proofs of these results can be seen in [DC01]. It is proved using the following remarkably result: For \( 1 \leq p < \infty \), **Q1** in \( L^p[-\pi, \pi] \) is equivalent to finding a constant \( C_p \), independent of \( k \) such that:

\[ \|S_k f\|_p \leq C_p \|f\|_p. \]

In that way, for \( p = 2 \), this is a consequence of Bessel Inequality [Cas+12, p. 99].
convergence almost everywhere, since we can modify a function almost everywhere without modifying the Fourier coefficients. But can we seek to obtain convergence everywhere if the function is a nice one, that is, at least continuous?

Working in nice spaces with some kind of regularity, the answer of Q2 is kind of positive. Thanks to Dini’s Criterion (see for example [DC01, p. 3]) for $2\pi$ periodic, continuous and bounded variation functions, their trigonometric series converge pointwise to the function everywhere. It is important to remark that the convergence of the trigonometric series at a point $x$ is a local property, that is, it depends on the value of the function in a neighbourhood of $x$; even though a modification of the function outside this neighbourhood may change the Fourier coefficients, it will never modify the convergence at the point $x$.

A natural following good space is the $2\pi$ periodic continuous functions, $C(\mathbb{T})$. In this space the answer is negative because the trigonometric series of a continuous function does not need to converge to the function. But indeed, if at a point does not converge to the function, the series must diverge! That is why the analysts of the $20^{th}$ century were interested about the convergence of the trigonometric series.

Du Boys Raymond was the first to give an explicit example of a continuous function whose trigonometric series diverges at one point. This fact can be proved theoretically (not constructive) using the Banach-Steinhaus Theorem (also known as Uniform Boundedness Principle)$^8$. It shows that the family of continuous functions whose Fourier series converges at a given $x$ is of first Baire category, in the Banach space of continuous functions on the circle.

So in some sense pointwise convergence is "atypical", and for most continuous functions the Fourier series does not converge at a given point. However, Carleson’s Theorem shows that, for a given continuous function, the Fourier series converges almost everywhere.

In general, the most common criteria for pointwise convergence of a $2\pi$ periodic continuous function $f$ are as follows:

- If $f$ satisfies a Hölder condition, then its Fourier series converges uniformly.
- If $f$ is of bounded variation, then its Fourier series converges everywhere.

$^8$A proof of this fact can be seen in [DC01, p. 5] and a proof of this theorem in [Cas+12, p. 264].
If its Fourier coefficients are absolutely summable, then the Fourier series converges uniformly.

Furthermore, Katznelson in [Kat66] showed that for any set of measure 0 there is a continuous periodic function whose Fourier series diverges at all points of the set (and possibly elsewhere). Katznelson, in the previous pages of this article, proved that the same assertion is even true for the space of functions $L^p[-\pi, \pi]$ with $1 \leq p < \infty$.

Combining Carleson’s theorem with Katznelson result shows that there is a continuous function whose Fourier series diverges at all points of a given set if and only if the set has measure 0.

To sum up, the important reformulation of $Q2$ is the convergence pointwise almost everywhere of the trigonometric series of functions belonging to spaces $C(\mathbb{T})$ and $L^p[-\pi, \pi]$ with $1 \leq p < \infty$.

The first positive answer in the positive direction of this reformulated question came in 1913 by Luzin [Luz13]. In this article he proved $Q2$ for functions in $L^2[-\pi, \pi]$ if the following conjecture, known as Luzin’s Conjecture, is true:

$$\lim_{n \to \infty} V.P. \int_{-\pi}^{\pi} \hat{f}(x-t) \frac{\cos(nt)}{t} dt = 0 \quad \text{a.e. in } x$$

And the first negative answer came ten years later, in 1923, by Kolmogorov. He gave an example of a function in $L^1[-\pi, \pi]$ which diverges almost everywhere (improved in 1926 to diverging everywhere).

In fact, the possible answer to this reformulation of $Q2$ in those spaces is one of the following options:

(+) Every trigonometric series converge almost everywhere.

(-) There exists a trigonometric series that diverges everywhere.

proved by Katznelson in [Kat66].

Before Carleson’s result, the best known estimate for the partial sums $S_n$ of the Fourier series of a function in $L^p[-\pi, \pi]$ was

$$S_n(x) = o((\log(n))^{1/p}) \text{ almost everywhere,}$$

This result had not been improved for several decades, leading some experts to suspect that it was the best possible and that Luzin’s conjecture was false. Kolmogorov’s counterexample in $L^1[-\pi, \pi]$ was unbounded in any interval, but it was thought to be only a matter of time before a continuous counterexample was found.

Carleson said in an interview with Raussen and Skau[RS07] that he started by trying to find a continuous counterexample and at one point thought he had a method that would construct one, but realized eventually that his approach could not work. He then tried instead to prove Luzin’s conjecture, since the failure of his counterexample convinced him that it was probably true.

In 1966, Carleson proved Luzin’s Conjecture [Car66].

Carleson’s original proof is exceptionally hard to read, and although several authors have simplified the argument there are still no easy proofs of his theorem. Expositions of the original paper of Carleson[Car66] include Kahane [Kah95], and Arias de Reyna [Ari02]. The book Grafakos [Gra09] also gives proofs of Carleson’s Theorem.

The extension of Carleson’s Theorem to $L^p$ for $p > 1$ was stated to be a rather obvious extension of the case $p = 2$ in Carleson’s paper, and was proved by Hunt [Hun68].

Charles Fefferman [Fef73] published a new proof of Hunt’s extension which proceeded by bounding a maximal operator. This, in turn, inspired a much simplified proof of the $L^2[-\pi, \pi]$, result by Michael Lacey and Christoph Thiele [LT00], explained in more detail in Lacey [Lac03].

That is why the main article we are going to follow in order to explain Carleson’s Theorem is [Lac03]. The techniques used are very important because they are related to time-frequency analysis, which has been very helpful recently in getting some results like, for example, about the Hilbert bilinear transform. This approach was also followed before by Juan Cavero de Carondelet [Car12]. His work has been very helpful in the redaction of these pages.
Extension of Carleson’s Theorem

Carleson’s result was improved further by Sjölin [Sjö71] to the space $L \log^+(L) \log^+ \log^+(L)^9$ and by Antonov to the space $L \log^+(L) \log^+ \log^+ \log^+(L)$.

Konyagin [Kon00] improved Kolmogorov’s counterexample by finding functions with everywhere-divergent Fourier series in a space slightly larger than $L \log^+(L)^{1/2}$. One can ask if there is in some sense a largest natural space of functions whose Fourier series converge almost everywhere. The simplest candidate for such a space that is consistent with the results of Antonov and Konyagin is $L \log^+(L)$.

The extension of Carleson’s Theorem to Fourier series and integrals in several variables is more complicated, as there are many different ways in which one can sum the coefficients; for example, one can sum over increasing balls, or increasing rectangles. Convergence of rectangular partial sums (and indeed general polygonal partial sums) follows from the one-dimensional case, but the spherical summation problem is still open for $L^2(\mathbb{R}^2)$.

In fact, the extension of question Q1 about the convergence in norm $p$ is not clear in several variables. Summing over polygonals, the result is the same as the unidimensional one and it is a consequence of it. But, for example, summing using the ball of the norm $L^2(\mathbb{R}^2)$, it is only true in the trivial case when $p = 2$. This was a problem of remarkable importance, known as the multiplier problem, which was solved by Fefferman [Fef71].

---

9Here $\log^+(L)$ is $\log(L)$ if $L > 1$ and 0 otherwise, and if $\phi$ is a function then $\phi(L)$ stands for the space of functions $f$ such that $\phi(f(x))$ is integrable.
Sketch of the proof

The proof of Carleson’s Theorem is divided in two chapters:

In the first chapter, we are going to reduce Carleson’s theorem progressively to different propositions:

• Using multiplier transference we are going to deal with its continuous equivalent reformulation. And to prove Carleson’s Theorem it is enough to see that a certain operator (Carleson operator) is bounded between $L^2(\mathbb{R})$ and Weak $L^2(\mathbb{R})$.

• In the spaces with no regularity it is very important to know the isometries of the space. So some results will be showed about this matter. This will help to understand better the Carleson operator.

• Later, we will explain what is the dyadic decomposition. This is the way that the time-frequency analysis appear in the proof.

• Then, with the language of the time-frequency analysis, using discretization we reduce the problem to the boundedness of a linear operator instead of a sublinear operator related to Carleson operator.

• Finally, our last step is to do a last reduction to prove the boundedness of this linear operator.

In the second chapter, we are going to prove this last reduction and consequently Carleson’s theorem. In order to do that:

• We need to explain some concepts related to the trees we get in the time-frequency analysis.

• We first state three technique lemmas about the trees, whose proofs are written later so that it does not interfere with the main idea of the result.

• Finally, with the help of these lemmas we can prove Carleson’s Theorem.
Is there a simplification of the proof for better function spaces?

There is no believe that there exist a simplification for the proof of Carleson’s Theorem when we restrict the attention to the space of continuous functions, even if we only want convergence somewhere rather than almost everywhere.

We now realise that pointwise convergence questions of $S_n f(x)$ are closely related to the boundedness properties of the Carleson maximal operator $C$.

This is only a sublinear operator, even though we would like it to be a linear operator. Roughly speaking, if one can prove a non-trivial bound on $C f$ (and in particular keep it finite almost everywhere) for all $f$ in a function space (e.g. $L^p[-\pi, \pi]$), then it is a relative routine matter to demonstrate almost everywhere convergence in $L^p[-\pi, \pi]$; and conversely, if no such bound exists, then it is likely that (with perhaps a bit of nontrivial trickery) one can eventually cook up a counterexample in this function space for which one has pointwise convergence nowhere.

If one has a bit of regularity, e.g. $C^1(\mathbb{T})$, then bounding $C f$ is relatively straightforward, but in spaces with zero regularity (E.g. $L^p[-\pi, \pi]$ or $C[-\pi, \pi]$) it is much more difficult.

The key problem here is the modulation invariance of $C$: if one multiplies $f$ by a character $e^{2\pi ikx}$, this essentially does not change $C$. The spaces $L^p[-\pi, \pi]$ and $C[-\pi, \pi]$ are also modulation invariant, and this basically forces any proof of almost everywhere (or even somewhere) pointwise convergence in these spaces to also be modulation invariant, which rules out a lot of standard techniques and requires instead tools such as time-frequency analysis.

Notation

$\lesssim$ It means $\leq C$, where $C$ is a constant whose value we do not care and it does not depend on the accompanying variables.

$\gtrsim$ It means $\geq C$, where $C$ is a constant whose value we do not care and it does not depend on the accompanying variables.

$\asymp$ Both $\lesssim$ and $\gtrsim$. 
Further remarks

In the bound of the Carleson operator, we use the boundedness of the maximal operator in $L^p$ with $p > 1$. Consequently, it is intuitive that the convergence of the series of Fourier almost everywhere is true for function in $L^p[-\pi, \pi]$ with $p > 1$. And as we have seen before this is the case.

The form of the Dirichlet kernel or the Lusin conjecture points out the essential difficulties in establishing the theorem. That part of the kernel or the conjecture that is convolution with $\frac{1}{x}$ corresponds to a singular integral. This can be done with the techniques associated to the Calderón Zygmund theory. In addition, one must establish some uniform control for oscillatory term in $\sin Nt$ or $\cos Nt$, which falls outside of what is commonly considered to be part of the Calderón Zygmund theory.
As we have already said, during this chapter we are going to develop conditions that imply Carleson’s Theorem.

1.1 Multiplier transference

We remind that Carleson’s theorem asserts that for every function $F \in L^2(\mathbb{T})$ its Fourier partial sum converges almost everywhere to $F$. That is,

$$F(t) = \lim_{N \to \infty} S_N F(t) = \lim_{N \to \infty} \sum_{n=-N}^{N} \hat{F}(n)e^{int} \text{ a.e. in } t$$

We recall that the Dirichlet kernel for series is defined by:

$$D_N(t) = \sum_{n=-N}^{N} e^{int} \text{ for } t \in \mathbb{T}$$

This kernel is related to Fourier partial sums by the equality: $S_N F(t) = F \ast D_N(t)$. Consequently, Carleson’s theorem can be reformulated as:

$$F(t) = \lim_{N \to \infty} S_N F(t) = \lim_{N \to \infty} F \ast D_N(t) \text{ a.e. in } t$$

This fact is known true for a dense subset of $L^2(\mathbb{T})$ (for instance $C^\infty(\mathbb{T})$), but it is not enough to prove the theorem. For that purpose, we have to introduce a maximal operator known as Carleson operator:
**Definition 1.1.1 Carleson operator \((\mathbb{T})\)**

The Carleson operator (for \(\mathbb{T}\)) is defined by:

\[
CF(t) = \sup_N |F * D_N(t)| = \sup_N \left| \sum_{n=-N}^{N} \hat{F}(n)e^{int} \right|
\]

for every \(F \in L^2(\mathbb{T})\).

We are going to prove a sufficient condition of Carleson’s theorem using the Carleson operator \((\mathbb{T})\).

**Theorem 1.1.2** If the Carleson operator \((\mathbb{T})\) is weak type \((2,2)\), then Carleson’s theorem holds.

**Proof.** We define:

\[
LF(t) = \limsup_{N \to \infty} \left| F(t) - \sum_{n=-N}^{N} \hat{F}(n)e^{int} \right|
\]

It is enough to prove that for every \(\varepsilon > 0\) holds \(|\{LF > \varepsilon\}| \lesssim \varepsilon\).

Let \(G \in C^\infty(\mathbb{T})\) be so that \(\|F - G\|_2 \leq \varepsilon^2\).

\[
|LF(t)| \leq |F(t) - G(t)| + C(F - G)(t)
\]

\[
|\{LF > \varepsilon\}| \leq |\{F - G > \varepsilon/2\}| + |\{C(F - G) > \varepsilon/2\}|
\]

By the theorem hypothesis,

\[
|\{C(F - G) > \varepsilon/2\}| \lesssim \frac{4}{\varepsilon^2} \|F - G\|_2^2 \lesssim \varepsilon
\]

Using Chebyshev’s inequality,

\[
|\{F - G > \varepsilon/2\}| \leq \frac{4}{\varepsilon^2} \|F - G\|_2^2 \lesssim \varepsilon
\]

We can define analogously the Carleson operator for \(f \in L^2(\mathbb{R})\).
**Definition 1.1.3 Carleson operator \((\mathbb{R})\)**

The Carleson operator \((\mathbb{R})\) is defined by:

\[
Cf(t) = \sup_N |f * D_N(t)| = \sup_N \left| \int_{-N}^{N} \hat{f}(\xi) e^{i\xi t} d\xi \right|
\]

for \(f \in L^2(\mathbb{R})\), where \(D_N\) is the Dirichlet kernel for integrals.

Multiplier transference guarantees that the Carleson operator \((\mathbb{T})\) is weak type \((2,2)\) if, and only if, the Carleson operator \((\mathbb{R})\) is. We are going to focus only in one implication:

**Theorem 1.1.4** Carleson operator \((\mathbb{R})\) is weak type \((2,2)\) \(\Rightarrow\) Carleson operator \((\mathbb{T})\) also is.

So in order to prove Carleson’s theorem we just need to see that the Carleson operator \((\mathbb{R})\) is bounded between \(L^2(\mathbb{R})\) and Weak \(L^2(\mathbb{R})\).

**Corollary 1.1.5** If Carleson operator \((\mathbb{R})\) is weak type \((2,2)\), then Carleson’s theorem holds.
1.2 Isometries

In the previous section we have seen a sufficient condition of Carleson’s theorem. In order to prove it, we need to introduce some operators in $L^2(\mathbb{R})$, which we will later see that are isometries of $L^2(\mathbb{R})$.

**Definition 1.2.1 Translation operator**

For every $f \in L^2(\mathbb{R})$, the translation operator is defined by:

$$\text{Tr}_y f(x) = f(x - y) \text{ with } y \in \mathbb{R}$$

**Definition 1.2.2 Modulation operator**

For every $f \in L^2(\mathbb{R})$, the modulation operator is defined by:

$$\text{Mod}_\xi f(x) = e^{ix\xi} f(x) \text{ with } \xi \in \mathbb{R}$$

**Definition 1.2.3 Dilation operator**

For every $f \in L^2(\mathbb{R})$, the dilation operator is defined by:

$$\text{Dil}_\lambda^2 f(x) = \lambda^{-1/2} f(x/\lambda) \text{ with } \lambda > 0$$

The appearance of these operators is recurrent in Harmonic Analysis and one of the properties that makes them important is that they are isometries.

**Proposition 1.2.4** These operators are well-defined isometries of $L^2(\mathbb{R})$.

*Proof.* Linearity and well-defined are obvious properties. It is enough to prove that these operators preserve the norm. And we can check this using a change of variables. ■

**Proposition 1.2.5** These operators are also isometries in Weak $L^2(\mathbb{R})$.

*Proof.* For translations, it is obvious due to the invariance of the measure by translations.

For modulation, it is trivial due to $|f(x)| = |\text{Mod}_\xi f(x)|$.

For dilation, we need to prove both inequalities:

$$\left| \{x \in \mathbb{R} : |\text{Dil}_\lambda^2 f(x)| > \alpha \} \right| = \lambda \left| \{s \in \mathbb{R} : |f(s)| > \alpha \lambda^{1/2} \} \right| \leq \alpha^{-2} \|f\|_{2,\infty}^2$$
Thus, \( \| \text{Dil}_{\lambda}^2 f \|_{2,\infty} \leq \| f \|_{2,\infty} \).

\[
\left\{ x \in \mathbb{R} : |f(x)| > \alpha \right\} = \lambda^{-1} \left\{ s \in \mathbb{R} : |\text{Dil}_{\lambda}^2 f(s)| > \alpha \lambda^{-1/2} \right\} \leq \alpha^{-2} \| \text{Dil}_{\lambda}^2 f \|_{2,\infty}^2
\]

Consequently, \( \| f \|_{2,\infty} \leq \| \text{Dil}_{\lambda}^2 f \|_{2,\infty} \)

We remember the definition of Fourier transform in \( L^2(\mathbb{R}) \).

**Definition 1.2.6 Fourier transform**

For every \( f \in L^2(\mathbb{R}) \), the Fourier transform is defined for all \( \xi \in \mathbb{R} \) by:

\[
\hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} dx
\]

The Fourier transform in \( L^2(\mathbb{R}) \) is also an isometry due to Parseval’s Theorem.

The relation between the previous isometries and the Fourier transform are shown in the following result:

**Proposition 1.2.7** These isometries are related with the Fourier transform in the following way, for every \( f \in L^2(\mathbb{R}) \):

\[
\begin{align*}
\text{Tr}_y f &= \text{Mod}_{-y} \hat{f} \\
\text{Mod}_\xi f &= \text{Tr}_\xi \hat{f} \\
\text{Dil}_{\lambda}^2 f &= \text{Dil}_{1/\lambda}^2 \hat{f}
\end{align*}
\]

**Proof.** Easy to proof with change of variables.

Now we are going to list some properties of these isometries that can be easily proved.

**Proposition 1.2.8** These isometries verify the following properties of composition:

- \( \text{Mod}_\xi \text{Mod}_\eta = \text{Mod}_{\xi+\eta} \)
- \( \text{Tr}_y \text{Tr}_z = \text{Tr}_{y+z} \)
- \( \text{Dil}_{\lambda}^2 \text{Dil}_{\mu}^2 = \text{Dil}_{\lambda \mu}^2 \)
- \( \text{Dil}_{\lambda}^2 \text{Mod}_{\xi} f(x) = \lambda^{-1/2} e^{i \xi x / \lambda} f(\frac{x}{\lambda}) = \text{Mod}_{\xi/\lambda} \text{Dil}_{\lambda}^2 f(x) \)
- \( \text{Dil}_{\lambda}^2 \text{Tr}_y f(x) = \lambda^{-1/2} f(\frac{x}{\lambda} - y) = \text{Tr}_{\lambda y} \text{Dil}_{\lambda}^2 f(x) \)
• $\text{Mod}_\xi \text{Tr}_y f = e^{i\xi} e^{i(x-y)\xi} f(x-y) = e^{i\xi} \text{Tr}_y \text{Dil}_2 f(x)$

And the following properties of adjointness, for every $f, g \in L^2(\mathbb{R})$:

• $\langle f | \text{Tr}_y g \rangle = \langle \text{Tr}_y f | g \rangle$
• $\langle f | \text{Mod}_\xi g \rangle = \langle \text{Mod}_\xi f | g \rangle$
• $\langle f | \text{Dil}_2 g \rangle = \langle \text{Dil}_2^2 f | g \rangle$

We are going to introduce a new operator which is related to the Carleson operator and has a very important property related to the isometries exposed in the next proposition.

**Definition 1.2.9 Operator $P_-$**

We define the operator $P_-$ for functions $f \in L^2(\mathbb{R})$ by

$$P_- f(x) = \int_{-\infty}^{0} \hat{f}(\xi)e^{ix\xi} d\xi$$

The operator $P_-$ plays a key rule in order to prove that the Carleson operator $(\mathbb{R})$ is weak type $(2,2)$ because:

$$\left| \int_{-N}^{N} \hat{f}(\xi)e^{ix\xi} d\xi \right| \leq \left| \int_{-\infty}^{N} \hat{f}(\xi)e^{i\xi t} d\xi \right| + \left| \int_{-\infty}^{-N} \hat{f}(\xi)e^{i\xi t} d\xi \right|$$

Hence, $Cf \leq C_1 f + C_2 f$ with:

$$C_1 f(x) = \sup_N \left| \int_{-\infty}^{N} \hat{f}(\xi)e^{i\xi t} d\xi \right|$$

$$C_2 f(x) = \sup_N \left| \int_{-\infty}^{-N} \hat{f}(\xi)e^{i\xi t} d\xi \right|$$

We can see the relation between $C_1$ with $P_-$ using $s = \xi + N$:

$$P_- \text{Mod}_N f(x) = \int_{-\infty}^{0} \hat{f}(\xi+N)e^{ix\xi} d\xi = e^{-ixN} \int_{-\infty}^{N} \hat{f}(s)e^{ixs} ds$$

Since $|e^{-ixN}| = 1$, we have proved that:

$$C_1 f(x) = \sup_N \left| P_- \text{Mod}_N f(x) \right|$$

Analogously,

$$C_2 f(x) = \sup_N \left| P_- \text{Mod}_N f(x) \right|$$
Remark 1.2.10  To prove that $C$ is weak-type $(2,2)$, it is enough to prove that both $C_1$ and $C_2$ are weak-type $(2,2)$.

The proposition that characterizes the operator $P_-$ is the following one:

**Proposition 1.2.11** Except a multiplying constant, $P_-$ is the only bounded operator in $L^2(\mathbb{R})$ such that:
1. Commutes with translations.
2. Commutes with dilations.
3. His kernel contains the functions $f$ such that $\text{sop}(\hat{f}) \subseteq [0, +\infty)$

**Proof.** It is clear that $P_-$ verifies these three conditions. Let us see if that an operator $T$ verifies these conditions then it must be a $P_-$, up to a multiplying constant. Because it commutes with translations, $T$ is given by a convolution with respect to a distribution. Such operators are equivalently characterized in frequency variables by $\hat{Tf} = \tau \hat{f}$ for some bounded function $\tau$. Commuting with dilations implies that $\tau(\xi) = \tau(\xi/|\xi|)$ for all $\xi \neq 0$. A function $f$ belongs to the kernel of $T$ if, and only if, $\hat{f}$ is supported on the zero set of $\tau$. Thus, the third condition implies that $\tau$ is identically 0 on the positive real axis, and non-zero on the negative axis. Hence, $T$ must be a multiple of $P_-$. ■
1.3 Dyadic decomposition

In order to use time-frequency analysis it is very important to know the dyadic decomposition of the real line.

**Definition 1.3.1 Dyadic interval**

A dyadic interval is an interval of the form:

\[ [j2^k, (j+1)2^k) \text{ with } j, k \in \mathbb{Z} \]

The set of all dyadic intervals will be denoted by \( \mathcal{D} \).

The measure of an interval \( I \) will be denoted by \( |I| \) and its center by \( c(I) \).

We are going to work in the time-frequency plane. We will say that a rectangle \( s = I \times w \in \mathcal{D} \times \mathcal{D} \) is a tile if \( |I|\|w| = 1 \). To improve the notation, we will generally write \( s = Is \times ws \). The set of all tiles will be denoted by \( T \). For every \( w \in \mathcal{D} \) we will name \( w^- \) and \( w^+ \) their inferior and superior halves respectively.

Since Fourier inversion is true in the Schwartz space \( S(\mathbb{R}) \), we can choose a function \( \varphi \in S(\mathbb{R}) \) such that:

\[
1_{\left[-\frac{1}{8}, \frac{1}{8}\right]} \leq \hat{\varphi} \leq 1_{\left[-\frac{1}{8}, \frac{1}{8}\right]}
\]

This implies that the support of \( \hat{\varphi} \) is contained in \( \left[-\frac{1}{8}, \frac{1}{8}\right] \).

From this \( \varphi \), we will get new \( \varphi_s \) for every \( s \in \mathcal{T} \), defined by:

\[
\varphi_s(x) := \text{Mod}_{c(ws)} Tr_{c(Is)} \text{Dist}_{Is}^2 \varphi(x) = |Is|^{-1/2} \varphi \left( \frac{x-c(Is)}{|Is|} \right) e^{ic(w^-)x}
\]

Consequently,

\[
\hat{\varphi}_s(\xi) = \text{Tr}_{c(ws^-)} \text{Mod}_{-c(Is)} \text{Dist}_{|ws|^2} \hat{\varphi}(\xi) = |ws^-|^{-1/2} \hat{\varphi} \left( \frac{\xi-c(ws^-)}{|ws^-|} \right) e^{i(c(ws^-)-\xi)c(Is)}
\]

We can locate the support of \( \hat{\varphi}_s \) inside the interval \( ws^- \) because:

\[
\frac{\left| \varphi - c(ws^-) \right|}{|ws^-|} \leq \frac{1}{8} \text{ implies } c(ws^-) - \frac{|ws^-|}{8} \leq \xi \leq c(ws^-) + \frac{|ws^-|}{8}
\]

Actually, the support is strictly contained in \( ws^- \). Furthermore, if we use the equality \( c(ws) = c(ws^-) + \frac{|ws|}{4} \) we can handle the support of \( \hat{\varphi}_s \) as a function of \( c(ws) \).

The following proposition is important for the next chapter in order to prove the technical lemmas. Since the following proposition is a technical result and its proof...
1.3 Dyadic decomposition

does not give us any idea or concept of the proof, we redirect the reader to the article of Lacey to consult its proof.

**Proposition 1.3.2** Let \( a, b, \mu, \nu \in \mathbb{R}, M > 1 \), then:

\[
\int_{\mathbb{R}} \frac{2^{\mu}}{(1 + 2^{\mu}|x - a|)^M} \frac{2^{\nu}}{(1 + 2^{\nu}|x - b|)^M} dx \lesssim \frac{2^{\min(\mu, \nu)}}{(1 + 2^{\min(\mu, \nu)}|a - b|)^M}
\]

And a consequence of this result, we can prove the following:

**Proposition 1.3.3** Let \( s, s' \in \mathcal{T} \), then:

\[
|\langle \varphi_s | \varphi_{s'} \rangle| \lesssim \min\left(\frac{|I_s|}{|I_{s'}|}, \frac{|I_{s'}|}{|I_s|}\right)^{1/2} \frac{2^\mu}{\left(1 + \frac{|c(I_s) - c(I_{s'})|}{\max(|I_s|, |I_{s'}|)}\right)^M}
\]

**Proof.** Suppose \( |I_s| = 2^i \) and \( |I_{s'}| = 2^j \), Let \( a = c(I_s) \) and \( b = c(I_{s'}) \). By the definition of the \( \varphi_s \) and using that \( \varphi \in S(\mathbb{R}) \):

\[
|\{ \varphi_s | \varphi_{s'} \}| \lesssim 2^{i+j} \int_{\mathbb{R}} \frac{2^{\mu}}{(1 + 2^{\mu}|x - a|)^M} \frac{2^{\nu}}{(1 + 2^{\nu}|x - b|)^M} dx
\]

Now, using 1.3.2 with \( \mu = -i \) and \( \nu = -j \), we get the result. ■

And in a particular case we can get the following estimate:

**Proposition 1.3.4** Furthermore, if \( |I_s| \leq |I_{s'}| \), then:

\[
|\langle \varphi_s | \varphi_{s'} \rangle| \lesssim \left(\frac{|I_s|}{|I_{s'}|}\right)^{1/2} \int_{I_{s'}} \frac{|I_s|^{-1}}{\left(1 + \frac{|x - c(I_s)|}{|I_s|}\right)^M} dx
\]

**Proof.** In this case \( \min(|I_s|, |I_{s'}|) = |I_{s'}| \), so using the previous proposition we get:

\[
|\langle \varphi_s | \varphi_{s'} \rangle| \lesssim \left(\frac{|I_{s'}|}{|I_s|}\right)^{1/2} \frac{1}{\left(1 + \frac{|c(I_s) - c(I_{s'})|}{|I_s|}\right)^M}
\]

Let us focus in the denominator of the second fraction; using the triangular inequality and taking \( x \in I_{s'} \), we get:

\[
\frac{|c(I_s) - c(I_{s'})|}{|I_s|} \geq \frac{|x - c(I_s)|}{|I_s|} - \frac{|x - c(I_{s'})|}{|I_s|} \geq \frac{|x - c(I_s)|}{|I_s|} - \frac{1}{2} \frac{|I_{s'}|}{|I_s|}
\]
Chapter 1. Problem Reduction

Considering that \( |I_s| \leq |I_s'| \):

\[
\frac{1}{\left( 1 + \frac{|c(I_s) - c(I_{s'})|}{|I_s|} \right)^M} \leq \frac{1}{\left( \frac{1}{2} + \frac{|x - c(I_s)|}{|I_s|} \right)^M} \lesssim \frac{1}{\left( 1 + \frac{|x - c(I_s)|}{|I_s|} \right)^M}
\]

But \( x \in I_{s'} \) is arbitrary and thanks to the integral mean value theorem, we know that:

\[
\frac{1}{\left( 1 + \frac{|c(I_s) - c(I_{s'})|}{|I_s|} \right)^M} \lesssim \frac{1}{|I_s'|} \int_{I_{s'}} \frac{1}{\left( \frac{1}{2} + \frac{|x - c(I_{s'})|}{|I_{s'}|} \right)^M} dx
\]

Getting the desired result:

\[
|\langle \phi_s | \phi_{s'} \rangle| \lesssim \left( \frac{|I_s|}{|I_{s'}|} \right)^{1/2} \int_{I_{s'}} \left( \frac{1}{\left( \frac{1}{2} + \frac{|x - c(I_{s'})|}{|I_{s'}|} \right)^M} \right) dx
\]

\[\blacksquare\]
1.4 Discretization

Now we are interested in obtaining a discrete operator in the discrete dyadic decomposition. With that aim we define:

**Definition 1.4.1 Operator $Q_\xi$**

Given $\xi \in \mathbb{R}$, we define then operator $Q_\xi$ between $L^2(\mathbb{R})$ by the formula:

$$Q_\xi f(x) = \sum_{s \in \mathcal{T}} 1_{w^+_s}(\xi) \langle f | \varphi_s \rangle \varphi_s(x)$$

For every $\xi \in \mathbb{R}$, there exists a sequence of dyadic intervals so that $\xi$ belongs to the superior half of every one of them. Depending on the base 2 of the number $\xi$, the involved intervals will have different measures. We will call these intervals $w(n)$ and we will force them to verify $|w(n)| = 2^{-n}$. We define $\mathcal{T}(n)$ as:

$$\mathcal{T}(n) = \{ s \in \mathcal{T} : w_s = w(n), \xi \in w^+_s \}$$

It may happen that, for some $n$, $\mathcal{T}(n) = \emptyset$.

It is obvious that if $n \neq m$, then $\mathcal{T}(n) \cap \mathcal{T}(m) = \emptyset$.

We can rewrite our operator $Q_\xi$

$$Q_\xi f = \sum_{n \in \mathbb{Z}} Q_{(n)} f \quad \text{with} \quad Q_{(n)} f = \sum_{s \in \mathcal{T}(n)} \langle f | \varphi_s \rangle \varphi_s$$

(if there is nothing to sum, we will think that the sum is 0).

This kind of decomposition is important because it verifies the following orthogonality condition:

**Proposition 1.4.2** If $n \neq m$, then:

$$\langle Q_{(n)} f | Q_{(m)} f \rangle = 0$$

**Proof.** By construction, $\xi \in w(n)^+ \cap w(m)^+$, and by the properties of the dyadic intervals, $w(n)^- \cap w(m)^- = \emptyset$. Let us compute the scalar product

$$\langle Q_{(n)} f | Q_{(m)} f \rangle = \left\langle \sum_{s \in \mathcal{T}(n)} \langle f | \varphi_s \rangle \varphi_s \right\rangle \left\langle \sum_{r \in \mathcal{T}(m)} \langle f | \varphi_r \rangle \varphi_r \right\rangle$$

$$= \sum_{s \in \mathcal{T}(n)} \sum_{r \in \mathcal{T}(m)} \langle f | \varphi_s \rangle \langle \varphi_r | \varphi_s \rangle \langle f | \varphi_r \rangle.$$
Chapter 1. Problem Reduction

For every \( s \in \mathcal{T}(n) \) and every \( r \in \mathcal{T}(m) \), we use that \( \text{sop}(\hat{\varphi}_s) \subset w(n)^- \) and \( \text{sop}(\hat{\varphi}_r) \subset w(m)^- \) and that the Fourier transform is an isometry in the Hilbert space \( L^2(\mathbb{R}) \); hence, 
\[
\langle \varphi_s | \varphi_r \rangle = \langle \hat{\varphi}_s | \hat{\varphi}_r \rangle = 0,
\]
because \( w(n)^- \cap w(m)^- = \emptyset \).

\[\blacksquare\]

In fact, two operators \( Q(n) \) and \( Q(m) \) differ by composition with modulation operators and dilation operators in the following way:

**Proposition 1.4.3** For every \( n, m \), there exist \( \lambda, \mu \) such that:

\[
Q(m) = \text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 Q(n) \text{Dil}_{1/\lambda}^2 \text{Mod}_{-\mu}
\]

**Proof.** We fix \( f \in L^2(\mathbb{R}) \) and take \( \xi \in \mathbb{R} \) such that \( \xi \in w(n)^+ \cap w(m)^+ \). Using the definition of \( Q(n) \) and the properties of the isometries 1.2.8 we get:

\[
\text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 Q(n) \text{Dil}_{1/\lambda}^2 \text{Mod}_{-\mu} f = \sum_{s \in \mathcal{T}(n)} \langle \text{Dil}_{1/\lambda}^2 \text{Mod}_{-\mu} f | \varphi_s \rangle \text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 \varphi_s
\]

\[
= \sum_{s \in \mathcal{T}(n)} \langle f | \text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 \varphi_s \rangle \text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 \varphi_s
\]

Taking \( \lambda = |w(n)|/|w(m)| \) and \( \mu = c(w(m)^-) - c(w(n)^-) / \lambda \) and using more properties of 1.2.8 and the definition of \( \varphi_s \) we get:

\[
\text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 \varphi_s = \text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 \text{Mod}_{c(w^-)} \text{Tr}_{c(I_s)} \text{Dil}_{1/|I_s|}^2 \varphi
\]

\[
= \text{Mod}_{\mu + c(w^-)/\lambda} \text{Tr}_{c(I_s)} \text{Dil}_{1/|I_s|}^2 \varphi
\]

\[
= \text{Mod}_{c(w^-)} \text{Tr}_{c(I_s)} \text{Dil}_{1/|I_s|}^2 \varphi = \varphi_r
\]

Consequently, with \( I_r = \lambda I_s \) and \( w_r = w(m) \):

\[
\text{Mod}_{\mu} \text{Dil}_{1/\lambda}^2 Q(n) \text{Dil}_{1/\lambda}^2 \text{Mod}_{-\mu} f = \sum_{r \in \mathcal{T}(m)} \langle f | \varphi_r \rangle \varphi_r = Q(m) f
\]

\[\blacksquare\]

With the help of the previous proposition we can reduce the proof of the following proposition to the case of a certain \( n \). We do not consider the proof relevant to understand properly the Carleson’s theorem, because it is just a technical result, so we omit it. However, it can be found in the reference article [Lac03].
**Proposition 1.4.4** For every \( \xi \in \mathbb{R} \), the operators \( Q_{(n)} \) are uniformly bounded (that is the constant of boundedness is independent of \( n \) and \( \xi \)).

For proving the following result, we use strongly the proposition above.

**Proposition 1.4.5** For every \( \xi \in \mathbb{R} \), the operator \( Q_\xi \) is a bounded operator with bound independent of \( \xi \).

**Proof.** It follows from the previous proposition and the orthogonality of the operator \( Q_{(n)} \).

The behaviour of the operator \( Q_\xi \) with respect to the dyadic intervals is shown in the following proposition:

**Proposition 1.4.6** Let \( \xi \in \mathbb{R} \) and \( k \in \mathbb{Z} \), then:

\[
Q_\xi = \text{Dil}_2^{-k} Q_{\xi_2^{-k}} \text{Dil}_2^k
\]

**Proof.** Following the same ideas of the proof of the result 1.4.3, this proof is straightforward.

The objective we persuade is to get an expression of the Carleson operator \( C_1 \) through the operator \( P_- \). Hence, we need to check the properties that verifies \( Q_\xi \).

**Proposition 1.4.7** The operator \( Q_\xi \) is positively semidefinite and its kernel contains those functions \( f \) such that \( \text{sop}(\hat{f}) \subseteq [\xi, +\infty) \).

**Proof.** Let us see first that it is positively semidefinite. For this purpose let us take \( f \in L^2(\mathbb{R}) \)

\[
\langle Q_\xi f , f \rangle = \sum_{n \in \mathbb{Z}} \sum_{s \in \mathcal{F}(n)} \vert \langle f, \varphi_s \rangle \vert^2 = \sum_{n \in \mathbb{Z}} \sum_{s \in \mathcal{F}(n)} \vert \langle \hat{f}, \hat{\varphi}_s \rangle \vert^2
\]

This is a sum of non negative terms. Considering that \( \text{sop}(\hat{\varphi}_s) \) covers all \( \mathbb{R} \), the result is trivial.

Let us now prove that its kernel contains those functions \( f \) such that \( \text{sop}(\hat{f}) \subseteq [\xi, +\infty) \).

Let \( f \in L^2(\mathbb{R}) \) be such that \( \hat{f} \) is supported in \( [\xi, +\infty) \). Recall that

\[
Q_\xi f = \sum_{s \in \mathcal{F}} 1_{w_2^+}(\xi) \langle f, \varphi_s \rangle \varphi_s
\]
Let \( s \in \mathcal{T} \) be such that \( \xi \in w^+_s \) (if this is not the case, the term will be zero). Because 
\[
\text{sop}(\hat{\phi}_s) \cap w^+_s = \emptyset, \quad \langle f | \phi_s \rangle = \langle \hat{f} | \hat{\phi}_s \rangle = 0.
\]
Consequently, \( f \) belongs to the kernel of \( Q\xi \), finishing in this way the proof. 

And the operator that we are looking for, which is related to the operator \( P_- \) and the dyadic composition, will be the following:

**Definition 1.4.8 Operator \( Q \)**

We define the operator \( Q \), at first, for functions of the Schwartz space \( f \) as follows:

\[
Qf(x) = \lim_{Y \to \infty} \int_{B(Y)} \text{Dil}^2_{2^{-\lambda}} \text{Tr}_{-y} \text{Mod}_{-\xi} Q\xi \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}^2_{2^{-\lambda}} f(x) \mu(d\lambda, dy, d\xi)
\]

where \( B(Y) = [1, 2] \times [0, Y] \times [0, Y] \) and the measure \( \mu(d\lambda, dy, d\xi) \) means that the integral is divided by \( |B(Y)| = Y^2 \).

For functions \( f \) of the Schwartz space, we can prove that the limit exists and so it is well-defined. Later, we can prove that it is bounded for these functions in the Schwartz space. Consequently, we can extend it by denseness to the whole \( L^2(\mathbb{R}) \). Finally, this important proposition holds.

**Proposition 1.4.9** The operator \( Q \) verifies the following properties:

1. Commutes with translations.
2. Commutes with dilations.
3. His kernel contains the functions \( f \) such that \( \text{sop}(\hat{f}) \subseteq [0, +\infty) \)

The proof of this proposition and the statements said before it are clearly technical and we refer the reader to the original article for their proofs.

**Corollary 1.4.10** \( Q = kP_- \)

**Proof.** This is just a consequence of proposition 1.2.11, because the hypothesis are true thanks to the previous proposition. 

This corollary is very important in the following problem reduction.

We need to know better \( C_1 \) and \( C_2 \). The main difficulty is that we do not know anything about the supremum. Even if the supremums are attained, they can be reached
at different points. The first step would be thinking of a measurable function \( N \) and defining the following operator.

**Definition 1.4.11** Operator \( C_N \)

For a measurable function \( N \), we define the operator by:

\[
C_N f(x) = \sum_{s \in \mathcal{S}} \mathbb{1}_{w_i} (N(x)) \langle f \rangle \phi_s(x)
\]

This is a very important key of the proof emphasized of Carleson’s theorem, emphasized in Fefferman’s paper [Fef73]. We should linearise the supremum. That is, we consider a measurable map \( N : \mathbb{R} \to \mathbb{R} \), which specifies the value of \( N \) at which the supremum in \( C_1 \) occurs. Then, it suffices to bound the operator norm of the linear (not sublinear) operator \( P_{-\text{Mod}_N(x)} \).

**Proposition 1.4.12** For every measurable function \( N \), \( C_N \) is weak-type \((2, 2)\). Then, \( C_1 \) and \( C_2 \) also are. Consequently, Carleson’s theorem holds.

**Proof.** By 1.4.10 we get that \(|P_{-\text{Mod}_{-N}f}| \simeq |Q_{\text{Mod}_{-N}f}|\).

\[
|P_{-\text{Mod}_{-N}f}| \lesssim \lim_{Y \to \infty} \int_{B(Y)} |\text{Dil}_{2^{-\lambda}} \text{Tr}_{-y} \text{Mod}_{-\xi} Q_{\xi} \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}_{2^{\lambda}} \text{Mod}_{-N} f| \mu(d\lambda, dy, d\xi)
\]

\[
\lesssim \lim_{Y \to \infty} \int_{B(Y)} |\text{Dil}_{2^{-\lambda}} \text{Tr}_{-y} \text{Mod}_{-\xi} Q_{\xi} \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}_{2^{\lambda}} f| \mu(d\lambda, dy, d\xi)
\]

Now we take suprema in \( N \) in both sides. Then, we take norms and use the convexity of the norm \( \| \cdot \|_{2, \infty} \). Hence, we get controlled except for a constant \( \|C_1 f\|_{2, \infty} \) by:

\[
\lim_{Y \to \infty} \int_{B(Y)} \left\| \sup_{N} |\text{Dil}_{2^{-\lambda}} \text{Tr}_{-y} \text{Mod}_{-\xi} Q_{\xi} \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}_{2^{\lambda}} f| \right\|_{2, \infty} \mu(d\lambda, dy, d\xi)
\]

By 1.2.5 and the non dependence of \( N \) in the left side, we get:

\[
\text{Vert} C_1 f \lesssim \lim_{Y \to \infty} \int_{B(Y)} \left\| \sup_{N} |Q_{\xi} \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}_{2^{\lambda}} f| \right\|_{2, \infty} \mu(d\lambda, dy, d\xi)
\]

Since the integral is normalized, it is enough to prove that the integrand is uniformly bounded in \( \lambda \), \( y \) and \( \xi \). However, this is equivalent, renaming the parameters, to:

\[
\left\| \sup_{N} |Q_{\xi} \text{Mod}_{\xi} \text{Tr}_{y} \text{Dil}_{2^{\lambda}} f| \right\|_{2, \infty} \lesssim \|f\|_2
\]
uniformly in $\lambda$, $y$ and $\xi'$. With this renaming we have placed $N$ where we wanted. Now we can use the hypothesis:

$$\|C_N \text{Mod}_{\xi'} \text{Tr}_y \text{Dil}^2_{24} f\|_{2,\infty} \lesssim \|\text{Mod}_{\xi'} \text{Tr}_y \text{Dil}^2_{24} f\|_2 = \|f\|_2$$

Consequently, we have proved the first part. The second part with $C_2$ is analogous. ■
1.5 Last reduction

Now, we are going to focus in the operator $C_N$ instead of the traditional Carleson one. There is a more suitable reduction to prove Carleson’s theorem. We remember that we just need to prove that for every measurable function $N$, $C_N$ is weak-type $(2,2)$.

Proposition 1.5.1 For every subset of finite measure $E$, every measurable function $N$ and every function $f \in L^2(\mathbb{R})$, if

$$|\langle C_N f|1_E \rangle| \lesssim \|f\|_2 |E|^{1/2}$$

holds, then $C_N$ is weak-type $(2,2)$.

Proof. We use that the dual of $L^{2,1}(\mathbb{R})$ is $L^{2,\infty}(\mathbb{R})$. Hence, using the definition of the dual:

$$\|C_N f\|_{2,\infty} \approx \sup \{ |\langle C_N f|g \rangle| : \|g\|_{2,1} \leq 1 \}$$

Since the simple functions are dense in the unit ball, it is enough to prove the inequality for all simple functions $g$ with norm less or equal than 1. Let $g = \sum_{j=1}^N a_j 1_{E_j}$ with $\|g\|_{2,1}$ and $a_1 > a_2 > \cdots > a_N > 0$. We define $F_j = \bigcup_{i=1}^j E_i$. We can rewrite $g$ as $g = \sum_{j=1}^N (a_j - a_{j+1}) 1_{F_j}$. Thus, using the hypothesis:

$$|\langle C_N f|g \rangle| \leq \sum_{j=1}^N (a_j - a_{j+1}) |\langle C_N f|1_{F_j} \rangle| \lesssim \|f\|_2 \sum_{j=1}^N (a_j - a_{j+1}) |1_{F_j}|^{1/2}$$

Reordering the sum, we have $\sum_{j=1}^N (a_j |F_j|^{1/2} - |F_{j-1}|^{1/2})$, and this sum is a multiple of $\|g\|_{2,1}$, finishing the proof. ■

And finally, we are going to give the last reduction of Carleson’s theorem. But, first, we need to introduce: $\psi_s = (1 + N) \phi_s$. Then,

$$|\langle C_N f|1_E \rangle| \leq \sum_{s \in \mathcal{F}} |\langle f|\phi_s \rangle\langle \psi_s|1_E \rangle|$$

Therefore, it is enough to prove the following theorem that would imply Carleson’s theorem:
Chapter 1. Problem Reduction

Theorem 1.5.2 For every function $f$ with $\|f\|_2 = 1$, every finite measurable $E$, every measurable function $N$ and every finite subset of tiles $S (S \subseteq \mathcal{T})$:

$$\sum_{s \in S} |\langle f|\varphi_s \rangle \langle \psi_s|1_E \rangle| \lesssim |E|^{1/2}$$

where the constant does not depend on $S$ or $N$.

Due to homogeneity, we can assume $\|f\|_2 = 1$, and since the constant does not depend on $S$, the series will converge.
2. The Proof of Carleson’s Theorem

In this chapter we deal with the proof of the last theorem of the previous chapter which we have already seen implies Carleson’s theorem. It is important to remark that this proof was due to Lacey and Thiele [LT00][Lac03], inspired by some ideas of the previous proof of the Carleson’s Theorem due to Fefferman [Fef73].

2.1 Idea of the proof

There exists a partial order in the set of all the tiles, defined by \( s \leq s' \) whenever \( I_s \subset I_{s'} \) and \( w_{s'} \subset w_s \). By the dyadic decomposition, two tiles are comparable if, and only if, they are not disjoint.

**Definition 2.1.1 Tree**

We will say that a subset of tiles \( T \) is a tree if there exists a tile \( I_T \times w_T \), named the peak of the tree, such that for every \( s \in T \), \( s < I_T \times w_T \).

**Definition 2.1.2 Count**

Let \( S \) be a subset of tiles. We will say that \( \text{count}(S) < A \) if for every decomposition of \( S \) in trees \( T_j \) (\( S = \bigcup_j T_j \)) we get \( \sum_j |I_{T_j}| < A \).

For \( M \) large enough, we define

\[
\chi(x) = \frac{1}{(1+|x|)^M}
\]
Chapter 2. The Proof of Carleson’s Theorem

As we have already done with $\varphi_s$, we would like to make this function suitable for our dyadic decomposition of the plane. That is, for every $I \in \mathcal{D}$ we define:

\[ \chi_I(x) = \text{Tr}_{c(I)} \text{Dil}_{|I|} \chi(x) = \frac{|I|^{-1}}{\left(1 + \frac{|x-c(I)|}{|I|}\right)^M} \]

Definition 2.1.3  Density

We define the density of a tile $s$ and a subset of tiles $S$ by:

\[ \text{den}(s) = \frac{1}{|E|} \sup_{s \subset u} \int_{E \cap N^{-1}(w_u)} \chi_u(x) \, dx \]

\[ \text{den}(S) = \sup_{s \in S} \text{den}(s) \]

Remark 2.1.4 Observe the monotonicity of the density.

Remark 2.1.5 Observe that if $M$ is chosen big enough we can get that $\text{den}(s) \leq |E|^{-1}$, because $\|\chi_u\|_1 = \|\chi\|_1$. Hence, $\text{den}(S) \leq |E|^{-1}$.

Definition 2.1.6  $\pm$ Tree

A $\pm$ tree is a tree $T$ with peak such that: for every $s \in T$ holds $w_T^\pm \subseteq w_s^\pm$.

Definition 2.1.7  Size

We define the size of a subset of tiles $S$ by:

\[ \text{size}(S) = \sup \{ |I_T|^{-1/2} \triangle(T) : T \text{ is a } + \text{ tree} \} \]

where $\triangle(T)^2 = \sum_{s \in T} |\langle f, \varphi_s \rangle|^2$.

Observe the monotonicity of the size.
2.2 Technical Lemmas

In this section we show three technique lemmas related to the concepts already defined about the trees (density, size, tree). Their proofs are written later so that it does not interfere with the main idea of the result.

**Lemma 2.2.1 — Density Lemma.**

For every finite subset of tiles $S$, we can express $S$ as the disjoint union of $S_l \cup S_h$ where:

\[
\text{den}(S_l) \leq \frac{1}{4}\text{den}(S)
\]

\[
\text{count}(S_h) \lesssim \text{den}(S)^{-1}
\]

**Lemma 2.2.2 — Size Lemma.**

For every finite subset of tiles $S$, it can be expressed as the disjoint union of $S_s \cup S_l$ where:

\[
\text{size}(S_s) \leq \frac{1}{2}\text{size}(S)
\]

\[
\text{count}(S_l) \lesssim \text{size}(S)^{-2}
\]

**Lemma 2.2.3 — Tree Lemma.**

For every tree $T$, holds:

\[
\sum_{s \in T} |\langle f | \varphi_s \rangle \langle \psi_s | 1_E \rangle| \lesssim |I_T|\text{size}(T)\text{den}(T)|E|
\]
2.3 The proof of Carleson’s Theorem

In order to prove Carleson’s theorem we need to prove this auxiliary proposition.

**Proposition 2.3.1** For a finite subset of tiles $S$, there exists a sequence of disjoint subsets $S_{n_0}, S_{n_0-1}, \ldots, S_0, S_1, \ldots$ such that:

1. $S = \bigcup_{n \leq n_0} S_n$
2. $\text{size}(S_n) \leq 2^{n+1}$ for every $n \leq n_0$.
3. $\text{den}(S_n) \leq 2^{n+2}$ for every $n \leq n_0$.
4. $\text{size}(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^n$ for every $n \leq n_0$.
5. $\text{den}(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^{2n}$ for every $n \leq n_0$.
6. $\text{cant}(S_n) \lesssim 2^{-2n}$ for every $n \leq n_0$.

**Proof.** We construct the sequence by regressive induction. Since $S$ is finite we can find $n_0 \in \mathbb{N}$ big enough such that:

$$\text{size}(S) \leq 2^{n_0}$$

We can choose $M$ big enough in order to make

$$\text{den}(S) \leq 2^{2n_0}$$

Hence, with $S_{n_0} = \emptyset$ we can verify from (2) to (6) in the proposition.

Suppose we have already defined $S_{n_0}, S_{n_0-1}, \ldots, S_n$ and that every one of them verifies from (2) to (6). We are going to construct $S_{n-1}$ using two auxiliary sets $S^1_{n-1}$ and $S^2_{n-1}$. We start with $S^1_{n-1}$. We know:

$$\text{den}(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^{2n}$$

Then, there are two cases:

1. Case 1: $\text{den}(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^{2(n-1)}$. We take $S^1_{n-1} = \emptyset$
2. Case 2: Otherwise, we use the Density lemma to get a $S^1_{n-1}$ such that:

$$\text{den}(S - (S_{n_0} \cup \cdots \cup S_n \cup S^1_{n-1})) \leq 2^{2(n-1)}$$

$$\text{cant}(S^1_{n-1}) \lesssim \text{den}(S - (S_{n_0} \cup \cdots \cup S_n)^{-1}) < 2^{-2n-1}$$
Now, we construct $S_{n-1}^2$ in a similar way. We know:

$$\text{size}(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^n$$

So there are two cases:

1. Case 1: size$(S - (S_{n_0} \cup \cdots \cup S_n)) \leq 2^{(n-1)}$. We take $S_{n-1}^2 = \emptyset$
2. Case 2: Otherwise, we use the Size lemma to get a $S_{n-1}^2$ such that:

$$\text{size}(S - (S_{n_0} \cup \cdots \cup S_n) \cup S_{n-1}) \leq 2^{(n-1)}$$

$$\text{cant}(S_{n-1}) \lesssim \text{size}(S - (S_{n_0} \cup \cdots \cup S_n)) \lesssim 2^{-2n-1}$$

We define $S_{n-1} = S_{n-1}^1 \cup S_{n-1}^2$. From the construction, it is easy to see that hold from (4) to (6). Besides, since $S_{n-1} \subseteq S - (S_{n_0} \cup \cdots \cup S_n)$ it is trivial to see that hold (2) and (3).

This algorithm ends in finite steps due to the finiteness of $S$ and condition (1) also holds.

---

**Theorem 2.3.2 — Carleson’s Theorem.**

For every function $f$ with $\|f\|_2 = 1$, every finite measurable $E$, every measurable function $N$ and every finite subset of tiles $S (S \subseteq \mathcal{T})$:

$$\sum_{s \in S} \left| \langle f | \varphi_s \rangle \langle \psi_s | 1_E \rangle \right| \lesssim |E|^{1/2}$$

where the constant does not depend on $S$ or $N$.

**Proof.** Let $S$ be a finite subset of tiles. We construct the sequence of tiles $S_n$ with $n \leq n_0$ using 2.3.1.

Using the condition 6 we know that $\text{count}(S_n) \lesssim 2^{-2n}$. Hence, by the definition of count, we know that there exists trees $T_j$ such that: $S_n = \cup_{j \in J_n} T_j$ and $\sum_{j \in J_n} |T_j| \lesssim 2^{-2n}$. Then,

$$\sum_{s \in S_n} \left| \langle f | \varphi_s \rangle \langle \psi_s | 1_E \rangle \right| \leq \sum_{j \in J_n} \sum_{s \in T_j} \left| \langle f | \varphi_s \rangle \langle \psi_s | 1_E \rangle \right|$$

$$\lesssim \sum_{j \in J_n} |T_j| \text{size}(T_j) \text{den}(T_j) |E|$$

$$\leq \sum_{j \in J_n} |T_j| \text{size}(S_n) \text{den}(S_n) |E|$$
where we have used the Tree Lemma in the second inequality, and the monotonicity of the size and density in the third one.

We have control of the size thanks to the proposition 2.3.1. This proposition also controls the size of $\sum_{j \in J_n} |I_T_j|$ by the definition of count. Consequently, $$\sum_{j \in J_n} |I_T_j| \text{size}(S_n) \text{den}(S_n) |E| \lesssim 2^{-2n} 2^{n+1} \text{den}(S_n) |E|$$

To control the density we have to use both proposition 2.3.1 and the remark of density. $$\text{den}(S_n) \leq \min(|E| - 1, 2^{2n+2})$$

Consequently, we get:

$$\sum_{s \in S_n} |\langle f, \varphi_s \rangle \langle \psi_s, 1_E \rangle| \lesssim 2^{1-n} \min(|E|^{-1}, 2^{2n+2}) |E|$$

$$\leq 2^{-n} \min(|E|^{-1}, 2^{2n}) |E|$$

$$= \min(2^n |E|^{-1/2}, 2^n |E|^{1/2}) |E|^{1/2}$$

Now we sum,

$$\sum_{s \in S_n} |\langle f, \varphi_s \rangle \langle \psi_s, 1_E \rangle| \lesssim \sum_{n \in \mathbb{Z}} \min(2^n |E|^{-1/2}, 2^n |E|^{1/2}) |E|^{1/2} \lesssim |E|^{1/2}$$

To finish the proof, we need to verify that the bound does not depend on $S$ or $|E|$.

$$\sum_{n \in \mathbb{Z}} \min(2^n |E|^{-1/2}, 2^n |E|^{1/2}) \leq \sum_{n \in \mathbb{Z}} 2^n |E|^{-1/2} + \sum_{n \in \mathbb{Z}} 2^n |E|^{1/2}$$

$$= 2 \left( |E|^{1/2} 2^{[\Omega]} + |E|^{-1/2} 2^{-[\Omega]} \right)$$

$$= 2 \left( 2^{[\Omega]} - \Omega + 2^{[\Omega] - \Omega} \right)$$

where we have defined $\Omega = \log_2(|E|^{-1/2})$. The last expression is bounded by a constant because $[\Omega] \leq \Omega < [\Omega] + 1$. ■
2.4 Proof of the Density lemma

Let $S$ be a finite subset of tiles. Let $\delta = \text{den}(S)$. We define:

$$S_h = \{ s \in S : \text{den}\{s\} \geq \frac{1}{4}\delta \}$$

Then, with $S_l = S - S_h$, it is trivial that: $\text{den}(S_l) \leq \frac{\delta}{4}$. Hence, to prove the lemma it is enough to see that:

$$\text{count}(S_h) \lesssim \delta^{-1}$$

Due to the definition of density, for all $s \in S_h$ exists a tile $u(s) \in \mathcal{T}$ with $u(s) > s$ such that:

$$\frac{1}{|E|} \int_{E \cap N^{-1}(w_u(s))} \chi_{I_u(s)}(x) dx \geq \frac{\delta}{4}$$

Let $\mathcal{U} = \{ u(s) : s \in S_h \}$ and $\mathcal{U}_{\text{max}}$ the subset of $\mathcal{U}$ with the maximal elements respect to the relation $<$. Let $S = \bigcup T_i$ where every $t_i = I_T \times w_{T_i}$ is a maximal tree in $S_h$. Given $t_i$, exists $u(t_i) \in \mathcal{U}$ such that $t_i < u(t_i)$. If $u(t_i) \notin \mathcal{U}_{\text{max}}$, then always exists $u \in \mathcal{U}_{\text{max}}$ such that $t_i < u(t_i) < u$. Anyway,

$$\sum_i |I_{T_i}| \leq \sum_{u \in \mathcal{U}_{\text{max}}} \sum_{t_i < u} |I_{T_i}|$$

On the other hand, for $u \in \mathcal{U}_{\text{max}}$ and $t_i \in T_i, t_j \in T_j$ such that $t_i < u$ and $t_j < u$. It is clear that $w_u \subseteq w_{T_i} \cap w_{T_j} \neq 0$ and since $t_i$ and $t_j$ are disjoint, it is necessary that $I_{T_i} \cap I_{T_j} = \emptyset$ and both intervals are contained in $I_u$. The previous inequality is now:

$$\sum_i |I_{T_i}| \leq \sum_{u \in \mathcal{U}_{\text{max}}} |I_u|$$

Consequently, it is left to prove the following:

$$\sum_{u \in \mathcal{U}_{\text{max}}} |I_u| \lesssim \delta^{-1}$$

For that purpose, we take $u \in \mathcal{U}_{\text{max}}$ and we express $\mathbb{R}$ as the disjoin union in $k \geq 0$ of the sets $(2^k I_u - 2^{k-1} I_u)$, where $2^k I_u$ is the interval with center $c(I_u)$ and length $2^k |I_u|$. Besides, we force that $2^{-1} I_u = \emptyset$. Thanks to this decomposition we get another expression from the density:

$$\frac{1}{|E|} \sum_{k=0}^{\infty} \int_{E_u} \chi_{I_u(s)}(x) dx \geq \frac{\delta}{8} \sum_{k=0}^{\infty} 2^{-k}$$
where $E_u = E \cap N^{-1}(w_{u(s)}) \cap (2^k I_u - 2^{k-1} I_u)$. Now we use the definition of $\chi_{I_{u(s)}}$ to get:

$$|E| \frac{\delta}{8} |I_u| \sum_{k=0}^{\infty} 2^{-k} < \sum_{k=0}^{\infty} \int_{E_a} \left(1 + \frac{|x-c(I_u)|}{|I_u|}\right)^M dx$$

Comparing term by term, we see that there exists a $k$ such a summand is less than the other. Furthermore, using that $x \in (2^k I_u - 2^{k-1} I_u)$, then $|x - c(I_u)| \geq 2^{k-2} |I_u|$ and we use this to bound the integral obtaining:

$$|I_u| < 82^k |E|^{-1} \delta^{-1} \frac{|E_u|}{(1 + 2^{k-2})^M}$$

Hence, $\mathcal{U}_{\text{max}} = \cup_k \mathcal{U}_k$ where

$$\mathcal{U}_k = \{ u \in \mathcal{U}_{\text{max}} : |I_u| \leq 8 \delta^{-1} |E|^{-1} 2^k (2^{k-2})^{-M} |E \cap N^{-1}(w_u) \cap 2^k I_u| \}$$

Consequently, it is enough to prove that:

$$\sum_{u \in \mathcal{U}_k} |I_u| \lesssim 2^{(2-M)k} \delta^{-1}$$

because:

$$\sum_{u \in \mathcal{U}_{\text{max}}} |I_u| \lesssim \sum_{k=0}^{\infty} 2^{(2-M)k} \delta^{-1} \lesssim \delta^{-1}$$

But in order to prove this sufficient condition we need to define recursively for every $n$ a subset of tiles $\mathcal{V}_k$ and some lemmas.

**Definition 2.4.1** $\mathcal{V}_k$

We start with $\mathcal{V}_k = \emptyset$. In the first step, we take $v_0 \in \mathcal{U}_k$ such that $|I_{v_0}|$ is the biggest as possible. We redefine $\mathcal{V}_k := \mathcal{V}_k \cup \{v_0\}$. For the second step, we take $v_1 \in \mathcal{U}_k - \{v_0\}$ such that $|I_{v_1}|$ is the biggest as possible with $[(2^k I_{v_1}) \times w_{v_1}] \cap [(2^k I_{v_0}) \times w_{v_0}] = \emptyset$. We redefine $\mathcal{V}_k := \mathcal{V}_k \cup \{v_1\}$. And we go on in the same way making $[(2^k I_{v_n}) \times w_{v_n}]$ disjoint from the previous ones and $|I_{v_n}|$ the biggest possible. This algorithm finish in a finite number of steps.

**Lemma 2.4.2** Given $u \in \mathcal{U}_k$, there exists $v \in \mathcal{V}_k$ such that: $[(2^k I_v) \times w_v] \cap [(2^k I_u) \times w_u] \neq \emptyset$ and $|I_u| \leq |I_v|$. We denote this situation $u \sim v$. 
Proof. If \( u \in \mathcal{Y}_k \). We take \( v = u \) and the lemma holds.

If not, \( u \) could not be add in the construction of \( \mathcal{Y}_k \) because \( [(2^k I_u) \times w_u] \) is not disjoint from some previous \( [(2^k I_{v_i}) \times w_{v_i}] \). Let \( i \) be the least index such that: \( [(2^k I_u) \times w_u] \cap [(2^k I_{v_i}) \times w_{v_i}] \neq \emptyset \). Because \( u \) was not chosen, we know that \( |I_u| \leq |I_{v_i}| \). We take \( v = v_i \).

\[ \]

**Lemma 2.4.3** Let \( v \in \mathcal{Y}_k \) and \( u, u' \in \mathcal{Y}_k \) (\( u \neq u' \)) such that: \( u \sim v \) and \( u' \sim v \). Then, \( I_u \cap I_{u'} = \emptyset \) and both are contained in \( 2^k I_v \).

*Proof.* Since \( u \) and \( u' \) are maximals respect to \( < \) and different we get \( u \cap u' = \emptyset \). Besides, since \( [(2^k I_v) \times w_v] \cap [(2^k I_u) \times w_u] \neq \emptyset \) and \( |I_u| \leq |I_v| \), we get \( w_v \subset w_u \). Analogously, \( w_v \subset w_{u'} \). Hence, \( w_v \subset w_u \cap w_{u'} \neq \emptyset \). Due to \( u \) and \( u' \) are disjoint we get: \( I_u \cap I_{u'} = \emptyset \).

The hypothesis:

\[ [(2^k I_v) \times w_v] \cap [(2^k I_u) \times w_u] \neq \emptyset \]

implies that \( 2^k I_u \subseteq 2^k I_v \). And analogously to \( u' \).

Using the two previous lemmas,

\[ \sum_{u \in \mathcal{Y}_k} |I_u| \leq \sum_{v \in \mathcal{Y}_k} \sum_{u \sim v} |I_u| \leq 2^k \sum_{v \in \mathcal{Y}_k} |I_v| \]

Furthermore, since every \( v \in \mathcal{Y}_k \) is also in \( \mathcal{Y}_k \), we get:

\[ \sum_{u \in \mathcal{Y}_k} |I_u| \leq 2^k \sum_{v \in \mathcal{Y}_k} 8 \delta^{-1}|E|^{-1} 2^{k-2} (2^k) M |E \cap N^{-1}(w_u) \cap 2^k I_u| \]

\[ \leq 2(2-M)k \delta^{-1} \sum_{v \in \mathcal{Y}_k} |E \cap N^{-1}(w_u) \cap 2^k I_u| / |E| \]

By construction, given \( v, v' \in \mathcal{Y}_k \) we have: \( [(2^k I_v) \times w_{v}] \cap [(2^k I_{v'}) \times w_{v'}] = \emptyset \).

So, \( (2^k I_v) \cap (2^k I_{v'}) = \emptyset \) or \( w_v \cap w_{v'} = \emptyset \). Anyway, \( E \cap N^{-1}(w_v) \cap 2^k I_v \) and \( E \cap N^{-1}(w_{v'}) \cap 2^k I_{v'} \) are disjoint and both contained in \( E \). Consequently,

\[ \sum_{u \in \mathcal{Y}_k} |I_u| \leq 2(2-M)k \delta^{-1} \]

proving the Density lemma.
2.5 Proof of the Size Lemma

Let $\mathcal{S}$ be a finite subset of tiles and let $\sigma = \text{size}(\mathcal{S})$. We are going to build a collection of disjoint trees $T \in \mathcal{T}_{\text{large}}$ such that

$$\mathcal{T}_{\text{large}} = \bigcup_{T \in \mathcal{T}_{\text{large}}} T \quad \text{and} \quad \sum_{T \in \mathcal{T}_{\text{large}}} |I_T| \lesssim \sigma^{-2}$$

For that, we will use a recursive procedure. Initialize

$$\mathcal{S}_{\text{stock}} := \mathcal{S}, \quad \mathcal{T}_{\text{large}} := \emptyset, \quad \mathcal{T}^+_{\text{large}} := \emptyset$$

We will repeat the process while $\text{size}(\mathcal{S}_{\text{stock}}) > \sigma/2$, as in the opposite case we will have $\text{size}(\mathcal{S}_{\text{stock}}) \leq \frac{1}{2} \text{size}(\mathcal{S})$ and will be enough to define $\mathcal{S}_{\text{small}} := \mathcal{S}_{\text{stock}}$ for stopping the procedure. We describe now this procedure. While $\text{size}(\mathcal{S}_{\text{stock}}) > \sigma/2$, we choose a $+\text{tree} T_+ \subseteq \mathcal{S}_{\text{stock}}$ such that $\Delta(T_+)/|I_{T_+}|^{-1/2} > \frac{\sigma^2}{2}$, i.e., such that

$$\sigma^2 |I_{T_+}| < 4\Delta(T_+)^2$$

In addition, the top of the tree $I_{T_+} \times \omega_{T_+}$ should be maximal with respect to the partial order $' \prec ' \text{among all trees that satisfy}$

$$\Delta(T_+) > \frac{\sigma^2}{2} |I_{T_+}|$$

And $c(\omega_{T_+})$ should be minimal, in the order of $\mathbb{R}$. Let $T$ now be the maximal tree in $\mathcal{S}_{\text{stock}}$ with top $I_{T_+} \times \omega_{T_+}$.

After this tree is chosen, update

$$\mathcal{S}_{\text{stock}} := \mathcal{S}_{\text{stock}} - T$$
$$\mathcal{T}_{\text{large}} := \mathcal{T}_{\text{large}} \cup \{T\}$$
$$\mathcal{T}^+_{\text{large}} := \mathcal{T}^+_{\text{large}} \cup \{T_+\}$$

When $\text{size}(\mathcal{S}_{\text{stock}}) \leq \frac{\sigma}{2}$, we set $\mathcal{S}_{\text{small}} := \mathcal{S}_{\text{stock}}$, and will be enough to check that $\text{count}(\mathcal{T}_{\text{large}}) \lesssim \text{size}(\mathcal{S})^{-2}$ holds, which is equivalent to the condition

$$\sum_{T \in \mathcal{T}_{\text{large}}} |I_T| \lesssim \sigma^{-2}$$

Therefore, it remains to verify this condition. But before doing that, we will enunciate and demonstrate a lemma which will allow us to see certain important properties of this construction.
2.5 Proof of the Size Lemma

**Lemma 2.5.1 — Strong disjointness.** Let $T^i_+$ and $T^j_+$ be two $+$-trees chosen by an inductive process after $i$ and $j$ steps, respectively. Let $s \in T^i_+$, $s' \in T^j_+$ be such that $\omega_s \subseteq \omega_{s'}$; then, $I_s \cap I_{T^i} = \emptyset$. Moreover, if $s' \in T^j_+$ and $s'' \in T^k_+$ are two different tiles such that there exists $s \in T^i_+$ with $\omega_s \subseteq \omega_{s'} \cap \omega_{s''}$, then $I_s \cap I_{s'} = \emptyset$.

**Proof.** By the definition of a $+$-tree, we know that $\omega_{T^j_+} \subseteq \omega_{s'}$. Moreover, by hypothesis, $\omega_s \subseteq \omega_{s'}$, so $\omega_{T^j_+} \subseteq \omega_{s'}$. In this situation, we have $c(\omega_{T^i_+}) < c(\omega_{T^j_+})$, and by minimality, necessarily $i < j$.

We want to check that $I_s \cap I_{T^i} = \emptyset$. Suppose the contrary. Then, taking into account that $T^i_+ = T^i$ and using the properties of dyadic intervals, we deduce that $I_s \subseteq I_{T^i}$. We obtain $s' < I_{T^i} \times \omega_{T^i}$, and by the maximality of $T^i$, we should have $s' \in T^i$, which was chosen before. We avoid this condition concluding that $I_s \cap I_{T^i} = \emptyset$.

For the second result, we will use what we have just proved. By hypothesis, for a certain $s \in T^i_+$ we have $\omega_s \subseteq \omega_{s'} \cap \omega_{s''}$, so we obtain three different cases.

If $\omega_s \subseteq \omega_{s''}$, we have previously seen that $I_s \cap I_{T^i} = \emptyset$, and, consequently, $I_{s'} \cap I_{s''} = \emptyset$. When $\omega_{s'} \subseteq \omega_{s''}$, the argument is analogous, so we only have to study the case $\omega_s = \omega_{s''}$. Taking into account that $s' \neq s''$ and $|I_{s'}| = |I_{s''}|$, we conclude that $I_{s'} \cap I_{s''} = \emptyset$.

We use this strong disjointness condition, and the selection criteria $\Delta(T^+_{T^i}) > \frac{\sigma}{2} |I_{T^i}|$ to prove the condition

$$\sum_{T \in R_{\text{large}}} |I_T| \lesssim \sigma^{-2}$$

For the election of $T^+$, we can write

$$\sigma^2 \sum_{T^+ \in R_{\text{large}}} |I_{T^+}| < \sum_{T^+ \in R_{\text{large}}} 4\Delta(T^+)^2 \lesssim \sum_{T^+ \in R_{\text{large}}} \sum_{s \in T^+} |\langle f | \phi_s \rangle|^2$$

Now, considering the self-adjoint operator $f \mapsto \langle f, \phi_s \rangle \phi_s$, if we define

$$\mathcal{I}^* := \bigcup_{T^+ \in R_{\text{large}}} \mathcal{T}$$

where we have already seen that all the $T^+$ are disjoint, and consider
Chapter 2. The Proof of Carleson’s Theorem

\[ F := \sum_{s \in S'} \langle f, \varphi_s \rangle \varphi_s, \]

then, we obtain

\[ \sigma^2 \sum_{T+ \in T_{\text{large}}} |I_{T+}| \lesssim \sum_{s \in S'} |\langle f|\varphi_s \rangle|^2 = \langle f|F \rangle \leq \| f \|_2 \| F \|_2 \]

Therefore, to conclude the proof will be enough to show

\[ \| F \|_2^2 \lesssim \sigma^2 \sum_{T+ \in T_{\text{large}}} |I_{T+}|, \]

because, in this case, we will have

\[ \sigma^2 \sum_{T \in T_{\text{large}}} |I_T| = \sum_{T+ \in T_{\text{large}}} |I_{T+}| \lesssim \sigma^{-2} \]

and this would conclude the proof of the Size Lemma. Hence, our main goal now is to prove the acotation

\[ \| F \|_2^2 \lesssim \sigma^2 \sum_{T+ \in T_{\text{large}}} |I_{T+}|, \]

For that, we begin expanding the term of the left side

\[ \| F \|_2^2 = \left\langle \sum_{s \in S'} \langle f|\varphi_s \rangle \varphi_s, \sum_{s \in S'} \langle f|\varphi_{s'} \rangle \varphi_{s'} \right\rangle \]

\[ \leq \sum_{s,s' \in S'} |\langle f|\varphi_s \rangle \langle f|\varphi_{s'} \rangle \langle \varphi_s|\varphi_{s'} \rangle| \]

\[ = \sum_{s,s' \in S'} |\langle f|\varphi_s \rangle \langle f|\varphi_{s'} \rangle \langle \varphi_s|\varphi_{s'} \rangle| + 2 \sum_{s,s' \in S'} |\langle f|\varphi_s \rangle \langle f|\varphi_{s'} \rangle \langle \varphi_s|\varphi_{s'} \rangle| \]

In the last inequality we have used that if \( \omega_s \cap \omega_{s'} = \emptyset \), then \( \langle \varphi_s|\varphi_{s'} \rangle = 0 \). Now, we will bound both of the terms of the right side separately.

**Lemma 2.5.2** The first term is bounded by

\[ \sum_{s,s' \in S', \omega_s = \omega_{s'}} |\langle f|\varphi_s \rangle \langle f|\varphi_{s'} \rangle \langle \varphi_s|\varphi_{s'} \rangle| \lesssim \sigma^2 \sum_{T+ \in T_{\text{large}}} |I_{T+}| \]

**Proof.** If we ordenate properly the terms of the previous sum, we can bound
\[
\sum_{s, s' \in S', \omega_s = \omega_{s'}} |\langle f \mid \varphi_s \rangle \langle f \mid \varphi_{s'} \rangle \langle \varphi_s \mid \varphi_{s'} \rangle| \leq \sum_{s \in S'} |\langle f \mid \varphi_s \rangle|^2 \sum_{s, s' \in S', \omega_s = \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|^2
\]

Using Proposition 1.3.4, the last inequality is bounded by
\[
\sum_{s \in S'} |\langle f \mid \varphi_s \rangle|^2 \sum_{s, s' \in S', \omega_s = \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|^2 \int_{I_s} \left( \frac{|I_s|^{-1}}{1 + \frac{x-c(I_s)}{|I_s|}} \right)^{M} dx
\]

On the other hand, the elements \( s' \in S' \) verifying \( \omega_s = \omega_{s'} \) have disjoint \( I_s \). The sum of the integrals is then bounded by \( \| \text{Tr}_{\{I_s\}} \text{Diil}_{I_s} \chi \|_1 = \| \chi \|_1 \). Using that \( |I_s| = |I_{s'}| \), the initial expression is bounded up to a constant by
\[
\sum_{s \in S'} |\langle f \mid \varphi_s \rangle|^2 = \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \sum_{s \in T_+} |\langle f \mid \varphi_s \rangle|^2 = \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \Delta(T_+)^2
\]

And obviously, \( \text{size}(S_{\text{stock}}) \leq \sigma \), so \( |T_+|^{-1} \Delta(T_+)^2 \leq \sigma^2 \); hence,
\[
\sum_{s, s' \in S', \omega_s \subseteq \omega_{s'}} |\langle f \mid \varphi_s \rangle \langle f \mid \varphi_{s'} \rangle \langle \varphi_s \mid \varphi_{s'} \rangle| \lesssim \sigma^2 \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} |T_+|
\]

The boundness of the second term is a little bit more complicated and we will divide it into several steps. Let us begin ordering the terms
\[
\sum_{s, s' \in S', \omega_s \subseteq \omega_{s'}} |\langle f \mid \varphi_s \rangle \langle f \mid \varphi_{s'} \rangle \langle \varphi_s \mid \varphi_{s'} \rangle| = \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \sum_{s \in T_+} |\langle f \mid \varphi_s \rangle| \sum_{s' \in S', \omega_s \subseteq \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|
\]

Now we use Cauchy-Schwarz in each of the factors along with the definition of \( \Delta(T_+) \). Then, the previous expression is bounded by
\[
\sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \left( \sum_{s \in T_+} |\langle f \mid \varphi_s \rangle|^2 \right)^{1/2} \left( \sum_{s' \in S', \omega_s \subseteq \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|^2 \right)^{1/2} \leq \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \Delta(T_+) \left( \sum_{s \in T_+} \left( \sum_{s' \in S', \omega_s \subseteq \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|^2 \right)^{1/2} \right) \leq \sum_{T_+ \in \mathcal{F}_{\text{large}}^+} \sigma |T_+|^{1/2} \left( \sum_{s \in T_+} \left( \sum_{s' \in S', \omega_s \subseteq \omega_{s'}} |\langle \varphi_s \mid \varphi_{s'} \rangle|^2 \right)^{1/2} \right)
\]
where we have used size\( (T_+) \leq \sigma \), and, hence, \( \Delta(T_+)^2 \leq \sigma^2 |I_T| \). Suppose now that we have already proved

\[
\sum_{s \in T_+} \left( \sum_{s' \in T', \omega_s \subseteq \omega_{s'}} |\langle f|\varphi_{s'}\rangle \langle \varphi_s|\varphi_{s'}\rangle| \right)^2 \lesssim \sigma^2 |I_T|
\]

Then, putting everything together, thanks to this acotation and the previous lemma, we have

\[
\|F\|^2 \lesssim \sigma^2 \sum_{T_+ \in \mathcal{T}_{\text{large}}} |I_T|,
\]

which is the expression we were looking for, so the Size Lemma is proved.

For proving the previous inequality, it is important to realise that \( \omega_s \subseteq \omega_{s'} \). Also, if we consider the set \( \{s'\} \), it is clear that

\[
|I_{s'}|^{-1/2} |\langle f|\varphi_{s'}\rangle| = \text{size}(\{s'\}) \leq \sigma
\]

Thanks to these observations, if we call

\[
A := \sum_{s \in T_+} \left( \sum_{s' \in T', \omega_s \subseteq \omega_{s'}} |\langle f|\varphi_{s'}\rangle \langle \varphi_s|\varphi_{s'}\rangle| \right)^2
\]

\[
B := \sum_{s \in T_+} \left( \sum_{s' \in T', \omega_s \subseteq \omega_{s'}} |I_{s'}|^{1/2} |\langle \varphi_s|\varphi_{s'}\rangle| \right)^2
\]

we have

\[
A \leq \sigma^2 B
\]

Therefore, for proving the previous inequality, it is enough to see that \( B \lesssim |I_T| \).

As \( |I_{s'}| \leq |I_s| \), we can apply Proposition 1.3.4, which, along with the definition of \( \chi_{I_s} \), implies

\[
|\langle \varphi_s|\varphi_{s'}\rangle| \lesssim \left( \frac{|I_s|}{|I_{s'}|} \right)^{1/2} \int_{I_{s'}} \chi_{I_s}(x)dx
\]

Replacing this in the expression of the definition of \( B \), we obtain
\[ B \lesssim \sum_{s \in T_+} |I_s| \left( \sum_{s' \in T_0, s \lesssim s'} \int_{I_{s'}} \chi_{I_{s'}}(x) \, dx \right)^2 \]

We can use now the Strong Disjointness Lemma to see that the elements \( s' \in T_0 \) such that \( \omega_s \subseteq \omega_s' \) verify \( I_s' \cap I_{T_+} = \emptyset \) and are mutually disjoint. Therefore, for every \( s \) appearing in the sum, we have \( I_s' \subseteq I_{T_+} \), so

\[ B \lesssim \sum_{s \in T_+} |I_s| \left( \int_{I_{T_+}} \chi_{I_s}(x) \, dx \right)^2 \leq \sum_{s \in T_+} |I_s| \int_{I_{T_+}} \chi_{I_s}(x) \, dx \]

because for \( M \) big enough, the term inside the brackets is less than 1. All the \( s \in T_+ \) appearing in the sum satisfy \( |I_s| \leq |I_{T_+}| \), and, in fact, there exists \( k \in \mathbb{N} \cup \{0\} \) such that \( |I_s| = 2^{-k}|I_{T_+}| \). This last thing allows us to write

\[ B \lesssim \sum_{k=0}^{\infty} \sum_{|I_s|=2^{-k}|I_{T_+}|} |I_s| \int_{I_{T_+}} \chi_{I_s}(x) \, dx \]

\[ = \sum_{k=0}^{\infty} 2^k \frac{1}{|I_{T_+}|} \sum_{s \in T_+, |I_s|=2^{-k}|I_{T_+}|} |I_s| \int_{I_{T_+}} \frac{1}{\left(1 + \frac{|x-c(I_s)|}{|I_s|}\right)^M} \, dx \]

We will prove in Lemma 2.5.3 the following inequality

\[ |I_s| \int_{I_{T_+}} \frac{1}{\left(1 + \frac{|x-c(I_s)|}{|I_s|}\right)^M} \, dx \lesssim \int_{I_s} \int_{I_{T_+}} \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \, dy \, dx \]

Therefore, if we fixe \( k \), as the sets \( I_s \) are disjoint and are contained in \( I_{T_+} \), we obtain

\[ B \lesssim \sum_{k=0}^{\infty} 2^k \frac{1}{|I_{T_+}|} \sum_{s \in T_+, |I_s|=2^{-k}|I_{T_+}|} \int_{I_s} \int_{I_{T_+}} \frac{1}{\left(1 + \frac{|x-y|}{2^{-k}|I_{T_+}|}\right)^M} \, dy \, dx \]

\[ \leq \sum_{k=0}^{\infty} 2^k \frac{1}{|I_{T_+}|} \int_{I_{T_+}} \int_{I_{T_+}} \frac{1}{\left(1 + \frac{|x-y|}{2^{-k}|I_{T_+}|}\right)^M} \, dx \, dy \]

For bounding the double integral, we will make use of Lemma 2.5.4, which states that there exists a constant \( C > 0 \) such that for every interval \( J \) and every \( b > 0 \) holds

\[ \int_{J} \int_{J} \frac{1}{\left(1 + \frac{|x-y|}{b|J|}\right)^M} \, dx \, dy \leq Cb^2 |J|^2 \]
Therefore,

\[ B \lesssim \sum_{k=0}^{\infty} \frac{2^k}{|I_{T+}|} \int_{I_{T+}} \int_{I_{T+}} \frac{1}{\left(1 + \frac{|x-y|}{2^k |I_{T+}|}\right)^M} \,dx \,dy \]

\[ \lesssim \sum_{k=0}^{\infty} \frac{2^k}{|I_{T+}|} \left(2^{-k} |I_{T+}|\right)^2 \]

\[ \lesssim |I_{T+}| \]

So we have proved the Size Lemma.

**Lemma 2.5.3** The following bounding holds:

\[ |I_s| \int_{I_{T+}} \int_{I_{T+}} \frac{1}{\left(1 + \frac{|x-c(I_s)|}{|I_s|}\right)^M} \,dx \,dy \lesssim \int_{I_s} \int_{I_{T+}} \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \,dx \,dy \]

**Proof.** For proving this, we take \( y \in I_s \) and use the triangular property

\[ \frac{|x-c(I_s)|}{|I_s|} \geq \frac{|x-y|}{|I_s|} - \frac{|y-c(I_s)|}{|I_s|} \geq \frac{|x-y|}{|I_s|} - \frac{1}{2} \]

Then, we can write

\[ \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \leq \frac{1}{\left(1 + \frac{|x-c(I_s)|}{|I_s|}\right)^M} \lesssim \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \]

As \( y \) is arbitrary, we can use the Mean Value Theorem for Integrals, so

\[ \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \lesssim \frac{1}{|I_s|} \int_{I_s} \frac{1}{\left(1 + \frac{|x-y|}{|I_s|}\right)^M} \,dy \]

Integrating the previous expression for \( x \in I_{T+}^c \), we obtain the result.

**Lemma 2.5.4** There exists a constant \( C > 0 \) such that for every interval \( J \) and every \( b > 0 \) holds

\[ \int_J \int_J \frac{1}{\left(1 + \frac{|x-y|}{b|J|}\right)^M} \,dx \,dy \leq Cb^2 |J|^2 \]
2.5 Proof of the Size Lemma

**Proof.** We move $x \mapsto x - c(J), y \mapsto y - c(J)$ and make a change of variable. Then, we obtain the sets $I$ and $I^c$, where $I = J - c(J) = (-\frac{1}{2}\alpha, \frac{1}{2}\alpha)$ and $\alpha = |J|$. For symmetry, we only have to bound the expression

$$\int_{I} \int_{I^c} \frac{1}{(1 + \frac{|x-y|}{ab})^M} dx dy \leq C\alpha^2 b^2$$

In this situation, $|x-y| = x - y$. For every $y \in I$, making the change of variable $s = 1 + \frac{x-y}{ab}$, we can write one of the integrals like

$$\int_{\frac{1}{2}\alpha}^{\infty} \frac{1}{(1 + \frac{|x-y|}{ab})^M} dx = \alpha b \int_{1 + \frac{1}{2b}}^{\infty} \frac{1}{s^M} ds \lesssim \frac{\alpha b}{(1 + \frac{1}{2b} - \frac{y}{ab})^{M-1}}$$

We make another change of variable, this time setting $t = 1 + \frac{1}{2b} - \frac{y}{ab}$ to obtain

$$\int_{I} \int_{I^c} \frac{1}{(1 + \frac{|x-y|}{ab})^M} dx dy \lesssim \int_{I} \frac{\alpha b}{(1 + \frac{1}{2b} - \frac{y}{ab})^{M-1}} dy$$

$$= \alpha^2 b^2 \int_{1}^{1 + \frac{1}{2b}} \frac{1}{t^{M-1}} dt$$

$$\lesssim \alpha^2 b^2$$

\[\blacksquare\]
2.6 Proof of the Tree lemma

Given a tree $T$ we are looking for to get:

$$\sum_{s \in T} |\langle f \varphi_s \rangle \langle \psi_s \mathbb{1}_E \rangle| \lesssim |I_T| \text{size}(T) \text{den}(T) |E|$$

For every $s \in T$ there exists $\varepsilon_s \in \mathbb{C}$ of module 1 such that:

$$|\langle f \varphi_s \rangle \langle \psi_s \rangle| = \varepsilon_s \langle f \varphi_s \rangle \langle \psi_s \rangle$$

So we can bound the terms by

$$\sum_{s \in T} |\langle f \varphi_s \rangle \langle \psi_s \mathbb{1}_E \rangle| = \sum_{s \in T} \varepsilon_s \langle f \varphi_s \rangle \int_E \psi_s(x) dx$$

$$\leq \int_E \left| \sum_{s \in T} \varepsilon_s \langle f \varphi_s \rangle \psi_s(x) \right| dx$$

where we have used that $T$ is a finite set of tiles. To handle the last integral we define

$$J' = \{ J \text{ dyadic : } I_s \not\subseteq 3J \text{ for every } s \in T \}$$

It is clear that given $x \in \mathbb{R}$, there is $J \in J'$ small enough such that $x \in J$. If we put $J \subset J'$ the subset of the maximal elements respect to the inclusion, then the intervals of $J$ are a partition of $\mathbb{R}$. They are disjoint due to the properties of the dyadic intervals and the maximality. From this, we get:

$$\sum_{s \in T} |\langle f \varphi_s \rangle \langle \psi_s \mathbb{1}_E \rangle| \leq \sum_{J \in J} \int_{E \cap J} \left| \sum_{s \in T} \varepsilon_s \langle f \varphi_s \rangle \psi_s(x) \right| dx$$

We divide our sum in two depending on the length of $I_s$ and we use the triangular inequality:

$$\sum_{s \in T} |\langle f \varphi_s \rangle \langle \psi_s \mathbb{1}_E \rangle| \leq \sum_{1} + \sum_{2}$$

where we have defined

$$\sum_{1} = \sum_{J \in J} \int_{E \cap J} \left| \sum_{s \in T \atop |I_s| \leq 2|J|} \varepsilon_s \langle f \varphi_s \rangle \psi_s(x) \right| dx$$

$$\sum_{2} = \sum_{J \in J} \int_{E \cap J} \left| \sum_{s \in T \atop |I_s| > 2|J|} \varepsilon_s \langle f \varphi_s \rangle \psi_s(x) \right| dx$$
2.6 Proof of the Tree lemma

We are going to bound this two sums up to a multiplicative constant by $|I_T| \text{size}(T) \text{den}(T)|E|$

Let us start with $\sum_1$ because it is simpler. For every $s \in T$ the tile $s$ is a $+$ tree, so:

$$|I_s|^{-1/2} \langle f | \varphi_s \rangle = \text{size}(\{s\}) \leq \text{size}(T)$$

Thanks to this observation and the fact that $|\varepsilon_s| = 1$. we can write

$$\sum_1 \leq \sum_{J \in J} \sum_{s \in T} |f| |\varphi_s| \int_{E \cap J} |\psi_s(x)| dx$$

$$\leq \text{size}(T) \sum_{J \in J} \sum_{s \in T} |s|^{1/2} |\psi_s(x)| dx$$

Remember that $\psi_s = (1_{w_s} \circ N) \varphi_s$ where $\varphi_s = \text{Mod}_{\ell(w_s)} \text{Tr}_{\ell(s)} \text{Dil}_{\ell(s)} \varphi$ and $\varphi$ is a Schwarz function, so

$$\sum_1 \leq \text{size}(T) \sum_{J \in J} \sum_{s \in T} \int_{E \cap J |N^{-1}(w_s^+)} |s|^{1/2} |\varphi_s(x)| dx$$

$$\leq \text{size}(T) \sum_{J \in J} \sum_{s \in T} \int_{E \cap J |N^{-1}(w_s^+)} \left( \frac{1}{1 + \frac{|x - c(l_s)|}{|s|}} \right)^2 dx$$

Using that $J \cap E \cap N^{-1}(w_s^+)$ is contained in $E \cap N^{-1}(w_s^+)$, $\sum_1$ is bounded by

$$\text{size}(T) \sum_{J \in J} \sum_{s \in T} |I_s| \left( \sup_{x \in J} \left( \frac{1}{1 + \frac{|x - c(l_s)|}{|s|}} \right)^M \right) \int_{E \cap N^{-1}(w_s^+)} \chi_{I_s}(x) dx$$

where $\chi_{I_s}$ is already defined in the section idea of the proof.

By the definition of $\text{den}(T)$ we can bound the integral by $\text{den}(T)|E|$:

$$\sum_1 \leq \text{size}(T) \text{den}(T) \left[ \sum_{J \in J} \sum_{s \in T} |I_s| \left( \sup_{x \in J} \left( \frac{1}{1 + \frac{|x - c(l_s)|}{|s|}} \right)^M \right) \right] |E|$$

To end this bound of $\sum_1$ is enough to prove that

$$\sum_{J \in J} \sum_{s \in T} |I_s| \left( \sup_{x \in J} \left( \frac{1}{1 + \frac{|x - c(l_s)|}{|s|}} \right)^M \right) \lesssim |I_T|.$$
This estimation is quite technical and we refer the article to its proof.

Returning to the second sum, $\sum_2$, is by far the more technical and we also refer to the article to the proof.
A. Background

This chapter’s aim is to show some basic background in order to understand Carleson’s Theorem proof. The basic ideas are in [Gra08].

A.1 Review of Lebesgue Spaces

Many quantitative properties of functions are expressed in terms of their integrability to a power. For this reason it is desirable to acquire a good understanding of spaces of functions whose modulus to a power $p$ is integrable. These are called Lebesgue spaces and are denoted by $L^p$.

Although we are mainly concerned with $L^p$ subspaces of Euclidean spaces, we discuss $L^p$ spaces of arbitrary measure spaces, since they represent a useful general setting.

Let $X$ be a measure space and let $\mu$ be a positive measure on $X$.

\begin{definition} Lebesgue Spaces $L^p(X, \mu)$ \\
For $0 < p < \infty$, $L^p(X, \mu)$ denotes the set of all complex-valued $\mu$-measurable functions on $X$ whose modulus to the $p$th power is integrable. \\
$L^\infty(X, \mu)$ is the set of all complex-valued $\mu$-measurable functions on $X$ such that for some $C > 0$, the set $\{x \in X : |f(x)| > C\}$ has $\mu$-measure zero. \\
Two functions in $L^p(X, \mu)$ are considered equal if they are equal $\mu$-almost everywhere. \end{definition}
Chapter A. Background

A.2 Weak Lebesgue Spaces

**Definition A.2.1** For $f$ a measurable function on $X$, the *distribution function* of $f$ is the function $d_f$ defined on $[0, \infty)$ as follows:

$$d_f(\lambda) = \mu(\{x \in X : |f(x)| > \lambda\})$$

The distribution function $d_f$ provides information about the size of $f$ but not about the behaviour of $f$ itself near any given point. For instance, a function on $\mathbb{R}^n$ and each of its translates have the same distribution function. It follows from the definition that $d_f$ is a decreasing function of $\lambda$ (not necessarily strictly).

**Example A.2.2** We are going to compute the distribution function of a non-negative simple function:

$$f(x) = \sum_{i=1}^{n} a_j \mathbb{1}_{E_j}(x)$$

where the sets $E_j$ are pairwise disjoint and $a_1 > a_2 > \cdots > a_n > 0$. If $\lambda \geq a_1$, then clearly $d_f(\lambda) = 0$. However, if $a_2 \leq \lambda < a_1$, then $|f(x)| > \lambda$ precisely when $x \in E_1$, and in general, if $a_{j+1} \leq \lambda < a_j$, then $|f(x)| > \lambda$ precisely when $x \in E_1 \cup \cdots \cup E_j$. Setting

$$B_j = \sum_{i=1}^{j} \mu(E_i)$$

we have

$$d_f(\lambda) = \sum_{i=0}^{n} B_i \mathbb{1}_{[a_{j+1}, a_j]}(\lambda)$$

where $a_0 = \infty$ and $B_n = a_{n+1} = 0$. The figure A.2 illustrates this example.

![Figure A.1: Example of distribution of a simple function](image-url)
A.2 Weak Lebesgue Spaces

**Proposition A.2.3** Let $f$ and $g$ be measurable functions on $(X, \mu)$. Then for all $\alpha, \beta > 0$ we have

1. $|g| \leq |f|$ $\mu$-a.e. implies $d_g \leq d_f$;
2. $\forall c \in \mathbb{C} \setminus 0$, $d_{cf}(\lambda) = d_f(\lambda/|c|)$;
3. $d_{f+g}(\alpha + \beta) \leq d_f(\alpha) + d_g(\beta)$;
4. $d_{fg}(\alpha \beta) \leq d_f(\alpha) + d_g(\beta)$

**Proof.** Easy for the reader to check. ■

Knowledge of the distribution function $d_f$ provides sufficient information to evaluate the $L^p$ norm of a function $f$ precisely. We state and prove the following important description of the $L^p$ norm in terms of the distribution function.

**Proposition A.2.4** For $f \in L^p(X, \mu)$, $0 < p < \infty$, we have

$$\|f\|_{L^p}^p = p \int_0^\infty \lambda^{p-1} d_f(\lambda) d\lambda$$

**Proof.** Indeed we have,

$$p \int_0^\infty \lambda^{p-1} d_f(\lambda) d\lambda = p \int_0^\infty \int_X 1_{\{|x| > \lambda\}} d\mu(x) d\lambda =$$

$$\int_X \int_0^{\|f(x)\|} p \lambda^{p-1} d\lambda d\mu(x) = \int_X \|f(x)\|^p d\mu(x) = \|f\|_{L^p}^p$$

where we used Fubini’s Theorem\(^1\) in the equality between the first and second lines. ■

**Definition A.2.5** Weak Lebesgue Spaces

For $0 < p < \infty$, the space weak $L^p(X, \mu)$ is defined as the set of all $\mu$-measurable functions $f$ such that:

$$\|f\|_{L^p, \infty} := \inf\{C > 0 : d_f(\lambda) \leq \frac{C^p}{\lambda^p} \forall \lambda > 0\} = \sup\{\gamma d_f(\gamma)^{\frac{1}{p}} : \gamma > 0\}$$

is finite. The space weak $L^\infty(X, \mu)$ is by definition $L^\infty(X, \mu)$.

---

\(^1\)Fubini’s Theorem: Suppose $X$ and $Y$ are measure spaces, and suppose that $X \times Y$ is given the maximal product measure (which is the only product measure if $X$ and $Y$ are $\sigma$-finite). If $f(x, y)$ is $X \times Y$ integrable, meaning that it is measurable and $\int_{X \times Y} |f(x, y)| d(x, y) < \infty$, then $\int_X \left( \int_Y f(x, y) dy \right) dx = \int_Y \left( \int_X f(x, y) dx \right) dy = \int_{X \times Y} f(x, y) d(x, y)$.
It is easy to check the certainty of the equality between the supremum and the
infimum above in the definition.

**Proposition A.2.6** Considering two functions in weak $L^p(X,\mu)$ equal if they are
equal $\mu$- a.e., weak $L^p(X,\mu)$ is a quasinormed space with the quasinorm defined
above.

**Proof.** Using Proposition A.2.3 (2), we can easily show that:

$$\|kf\|_{L^p,\infty} = |k|\|f\|_{L^p,\infty}$$

Using Proposition A.2.3 (3), we get:

$$\|f + g\|_{L^p,\infty} \leq c_p(\|f\|_{L^p,\infty} + \|g\|_{L^p,\infty})$$

Actually, it can be proved that this space is complete. And considering the different
cases of the value of $p$ in the triangular inequality we get the following result:

**Proposition A.2.7** In fact, weak $L^p(X,\mu)$ with $p \geq 1$ is a Banach space. And for
$0 < p < 1$ a quasi-Banach space.

The importance of the weak Lebesgue spaces is to use the following definitions
that are vital for Carleson’s theorem proof.

**Definition A.2.8** type $(p,q)$
An operator $T$ between $L^p$ and $L^q$ is said to be of type $(p,q)$ if it is bounded.

**Definition A.2.9** type $(p,q)$
An operator $T$ between $L^p$ and weak $L^q$ is said to be of weak type $(p,q)$ if it is
bounded.
A.3 Maximal Operator

This concept of maximal operator is fundamental to know in order to understand the tree lemma.

**Definition A.3.1** Hardy–Littlewood maximal operator

The Hardy–Littlewood maximal operator $M$ is a significant non-linear operator used in real analysis and harmonic analysis. It takes a locally integrable function $f : \mathbb{R}^n \rightarrow \mathbb{C}$ and returns another function $Mf$ that, at each point $x \in \mathbb{R}^n$, gives the maximum average value that $f$ can have on balls centred at that point. More precisely,

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| \, dy$$

**Proposition A.3.2** Weak Type Estimate

For $n \geq 1$ and $f \in L^1(\mathbb{R}^n)$, there is a constant $C_n > 0$ such that for all $\lambda > 0$, we have:

$$|\{Mf > \lambda\}| < \frac{C_n}{\lambda} \|f\|_{L^1(\mathbb{R}^n)}.$$

With the Hardy–Littlewood maximal inequality in hand, the following strong-type estimate is an immediate consequence of the Marcinkiewicz interpolation theorem:

**Proposition A.3.3** Strong Type Estimate

For $n \geq 1$, $1 < p \leq \infty$, and $f \in L^p(\mathbb{R}^n)$, there is a constant $C_{p,n} > 0$ such that

$$\|Mf\|_{L^p(\mathbb{R}^n)} \leq C_{p,n} \|f\|_{L^p(\mathbb{R}^n)}.$$
Bibliography

Books


Articles


Chapter A. Background


## Index

### Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_k$</td>
<td>38</td>
</tr>
<tr>
<td>$\pm$ Tree</td>
<td>32</td>
</tr>
<tr>
<td>$\varphi_s$</td>
<td>20</td>
</tr>
</tbody>
</table>

### C

- Carleson’s Theorem | 1
- Reduction | 35
- Count | 31

### D

- Density | 32
- Lemma | 33
- Dirichlet kernel | 13
- Dyadic interval | 20

### F

- Fourier transform | 17

### H

- Hardy–Littlewood maximal operator | 55

### L

- Lebesgue Spaces | 51
  - Weak | 53

### O

- Operator
  - $C_N$ | 27
  - $P_-$ | 18
  - $Q_-$ | 26
  - $Q_{\xi}$ | 23
  - Carleson ($\mathbb{R}$) | 15
  - Carleson ($\mathbb{T}$) | 14
  - dilation | 16
  - modulation | 16
  - translation | 16
S

Size ........................................ 32
Lemma .................................... 33

T

Tree ........................................ 31
Lemma .................................... 33
type \((p,q)\) ................................. 54
weak .......................................... 54