Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Triennale in Matematica

The Spherical Maximal Function: Classical and Geometric Approaches

Relatore:
Prof. Paolo Ciatti

Laureando: Marco Locatelli
Matricola: 2122133

Anno Accademico 2025/2026

17/07/2026

Contents

Introduction

The study of maximal operators is rooted in the problem of convergence for averages of functions. In 1930, Hardy and Littlewood established foundational inequalities for the maximal operator bearing their name [4]. This result led to a proof of the Lebesgue Differentiation Theorem, making clear that the underlying analytical problem for the convergence of such averages is the control of appropriate maximal operators. Although the problem of averaging over a broader variety of sets received much attention in the early 1930s (see also [2]), it was not until the 1970s that Stein and Strömberg, in their attempt to find dimension-independent constants for the Hardy-Littlewood Maximal Inequality [13], stumbled upon the natural yet harder problem of averaging over lower-dimensional sets, such as spheres. In 1976, Stein proved the L^p boundedness of the Spherical Maximal Operator for $d \geq 3$, in the sharp range of exponents $\frac{d}{d-1} < p \leq \infty$ [9]. His argument, which relies heavily on Fourier analysis, exploits the decay properties of the Fourier transform of the spherical measure through Plancherel's theorem and his newly developed method of square functions. The planar case proved far more challenging. It was only ten years later, in 1986, that Bourgain resolved it in a highly complex and intricate work that blended the inherent geometric nature of the problem with Fourier analysis. In general, Fourier analytical tools are remarkably powerful in the theory of geometric maximal functions. Rubio de Francia [7] established a far-reaching extension of Stein's theorem for a class of maximal Fourier multiplier operators, which allows for the study of maximal averages over a large variety of surfaces. His work relies on a deep underlying relationship between the curvature of a surface and the decay of the Fourier transform of its surface measure. Despite the power of Fourier analytical tools, it is also of interest to understand maximal operators, such as the spherical one, from a purely geometric and combinatorial standpoint, without using the Fourier transform. An instance of this is Marstrand's circle packing theorem [6], which is a precursor to the circular maximal theorem. Also truly remarkable is a completely geometric proof of Bourgain's theorem given by W. Schlag [8]. The aim of the present work is to examine techniques arising from both approaches to the spherical maximal function. Classical methods will be discussed in the first two chapters, retracing Stein's proof of the spherical maximal theorem. In Chapter 3, we shift our perspective to a purely geometric one, highlighting a series of key reductions and ideas that will be used in the following part of the work. Finally, in Chapter 4, we will combine geometric intuition with Fourier analytic techniques to provide an alternative proof of Stein's theorem for the local version of the operator. As we will see, these ideas extend naturally to the plane and will shed further light on the intricate work of Bourgain, ultimately highlighting the true enemy to overcome: tangency.

Chapter 1

Harmonic Analysis Preliminaries

Before introducing the spherical maximal function, the central object of this thesis, it is convenient to collect some tools from Harmonic Analysis that will be used systematically in the proofs to come. The results stated in the first three sections can be found in [3], while the theory of oscillatory integrals and the Fourier transform of the spherical measure is drawn from [12] and [10].

1.1 Hardy-Littlewood Maximal Theorem

Definition 1.1. Let f be a locally integrable function. The *Hardy-Littlewood* or *standard maximal function* is

$$\mathcal{M}f(x) = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)|\, dy,$$

where the supremum is taken over the family of balls containing x.

Theorem 1.2 (Hardy-Littlewood)**.** *The following estimates hold:*

(i) *A weak $L^{1,\infty}$ bound:*

$$\|\mathcal{M}f\|_{L^{1,\infty}} \lesssim \|f\|_{L^1}.$$

(ii) *If $f \in L^p(\mathbb{R}^d)$ for $1 < p \leq \infty$, a strong L^p bound:*

$$\|\mathcal{M}f\|_{L^p} \lesssim \|f\|_{L^p}.$$

Where the implicit constant is independent of f for both.

1.2 Approximation to the identity

Definition 1.3. For any integrable function Φ we set

$$\Phi_t(x) = t^{-d}\Phi(x/t), \quad t > 0.$$

If $\int_{\mathbb{R}^d} \Phi = 1$, we call the family $\{\Phi_t\}_{t>0}$ an *approximation to the identity.*

Theorem 1.4. *Let Φ be a function on $\mathbb{R}^d$ that has an integrable radially decreasing majorant and such that $\int_{\mathbb{R}^d} \Phi = 1$. If $f \in L^p$ for some $p < \infty$, then*

$$\lim_{t\to 0} f * \Phi_t(x) = f(x)$$

for a.e. $x \in \mathbb{R}^d$.

Corollary 1.5. *Let Φ be a function on $\mathbb{R}^d$ that has an integrable radially decreasing majorant. Let $a = \int_{\mathbb{R}^d} \Phi(x)\, dx$. Then for all $f \in L^p(\mathbb{R}^d)$ and $1 \le p < \infty$,*

$$(f * \Phi_t)(x) \to a\, f(x)$$

for a.e. $x \in \mathbb{R}^d$ as $t \to 0$.

Lemma 1.6. *For any nonnegative, radial decreasing function $\Psi \in L^1$, we have*

$$|f * \Psi(x)| \le \mathcal{M} f(x) \int \Psi(y)\, dy.$$

1.3 Interpolation

Of crucial importance for the proof of the spherical maximal theorem is the use of interpolation between Lebesgue spaces for linear or sublinear operators. The two main results, are the following.

Theorem 1.7 (Marcinkiewicz)**.** *Let (X, μ) be a σ-finite measure space, let (Y, ν) be another measure space, and let $0 < p_0 < p_1 \le \infty$. Let T be a sublinear operator defined on $L^{p_0}(X) + L^{p_1}(X) = \{f_0 + f_1 : f_j \in L^{p_j}(X),\ j = 0, 1\}$ and taking values in the space of measurable functions on Y. Assume that there exist $A_0, A_1 < \infty$ such that*

$$\|T(f)\|_{L^{p_0,\infty}(Y)} \le A_0 \|f\|_{L^{p_0}(X)} \qquad \text{for all } f \in L^{p_0}(X),$$
$$\|T(f)\|_{L^{p_1,\infty}(Y)} \le A_1 \|f\|_{L^{p_1}(X)} \qquad \text{for all } f \in L^{p_1}(X).$$

Then for all $p_0 < p < p_1$ and for all $f \in L^p(X)$ we have the estimate

$$\|T(f)\|_{L^p(Y)} \le A \|f\|_{L^p(X)},$$

where

$$A = 2\left(\frac{p}{p - p_0} + \frac{p}{p_1 - p}\right)^{\frac{1}{p}} A_0^{\frac{\frac{1}{p} - \frac{1}{p_1}}{\frac{1}{p_0} - \frac{1}{p_1}}} A_1^{\frac{\frac{1}{p_0} - \frac{1}{p}}{\frac{1}{p_0} - \frac{1}{p_1}}}.$$

Theorem 1.8 (Riesz-Thorin)**.** *Let (X, μ), (Y, ν) be measure spaces. Let $p_0, p_1, q_0, q_1 \in [1, \infty]$, and $T : (L^{p_0} + L^{p_1})(X) \to (L^{q_0} + L^{q_1})(Y)$ be a linear operator. Suppose there exist constants A_0, A_1 such that*

$$\|Tf\|_{L^{q_0}} \le A_0 \|f\|_{L^{p_0}} \quad \text{for all } f \in L^{p_0}(X),$$
$$\|Tf\|_{L^{q_1}} \le A_1 \|f\|_{L^{p_1}} \quad \text{for all } f \in L^{p_1}(X).$$

Then for any $\theta \in (0, 1)$, we have

$$\|Tf\|_{L^q} \le A_0^{1-\theta} A_1^{\theta} \|f\|_{L^p}$$

for $f \in L^p(X)$, where

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

1.4 Uncertainty Principle

In its many formulations, the Uncertainty Principle dictates much of the geometry of the Fourier Transform, for our purposes we will consider the following.

> [**Uncertainty Principle:**] "If f is spatially concentrated at scale R, then $\hat{f}$ is locally constant at scale $\frac{1}{R}$".

We now give two examples illustrating this heuristic.
Example 1: For $R \geq 1$ let

$$f_R(x) = e^{\frac{-|x|^2}{R^2}}$$

be the Gaussian with "variance" R. f_R is spatially concentrated on $B(0,R)$, since for any $\varepsilon > 0$ if

$$|x| \geq R^{1+\varepsilon}, \quad \text{then} \quad |f_R(x)| \leq C_{N,\varepsilon} R^{-N} \quad \text{for all } N \in \mathbb{N}.$$

We know that $\hat{f}(0) = \int_{\mathbb{R}^d} f = \pi^{d/2} R^d$ so, by the uncertainty principle, we expect

$$\hat{f}_R(\xi) \approx \pi^{d/2} R^d \quad \text{for } |\xi| \leq R^{-1}.$$

Indeed, from the well-known formula

$$\hat{f}_R(\xi) = \pi^{d/2} R^d e^{-\pi^2 R^2 |x|^2},$$

this is the case. Furthermore, for $\varepsilon > 0$ if

$$|\xi| \geq R^{-1+\varepsilon}, \quad \text{then} \quad |\hat{f}_R(\xi)| \leq C_{N,\varepsilon} R^{-N} \quad \text{for all } N \in \mathbb{N}.$$

So $\hat{f}_R$ is 'essentially' 0 on all other R^{-1} balls; thus, f_R respects the uncertainty principle.

Example 2: More generally, let $\phi \in \mathcal{S}(\mathbb{R}^d)$ be a Schwartz function and define

$$\phi_R(x) := \phi(R^{-1}x), \quad \text{then} \quad (\phi_R)^\wedge(\xi) = R^d \widehat{\phi}(R\xi).$$

It follows that

- ϕ_R is 'concentrated on $B(0,R)$
- $\widehat{\phi}_R$ is 'concentrated on $B(0,1/R)$

Moreover, $\widehat{\phi}_R$ is 'locally constant' at scale $1/R$. Indeed, by the chain rule, we have $\nabla\widehat{\phi}_R(\xi) = R^{d+1}\nabla\widehat{\phi}(R\xi)$. On a ball of radius $1/R$, the maximum variation of $\widehat{\phi}_R$ is roughly bounded by $R^{d+1} \cdot R^{-1} = R^d$, which is proportional to its peak amplitude. This means that the function remains stable on this scale.

1.5 Oscillatory Integrals

Integrals of the form

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\Phi(x)}\psi(x)\,dx,$$

appear almost everywhere in harmonic analytic contexts, and they go under the label of *Oscillatory Integrals*. Usually, the function Φ is called the *phase*, while ψ is the *amplitude*. The theory of oscillatory integrals is very broad and complex; for our purposes, we will delve only into the basic ideas and results needed for the proofs coming up later on. We will assume $\lambda > 0$, Φ and ψ smooth, Φ real-valued and ψ compactly supported.

Stationary Phase Principle

The general principle of the theory is that of "Stationary Phase". The idea underlying the principle, as we will see, is that the main contributions to $I(\lambda)$ come from points where the phase is stationary ($\nabla\Phi = 0$), while in the regions where the gradient is non-vanishing the integral decays quickly as $\lambda \to \infty$.

Lemma 1.9. *Suppose $|\nabla\Phi(x)| \geq c > 0$ for all x in the support of ψ. Then*

$$|I(\lambda)| \lesssim \lambda^{-N}$$

for all $N \in \mathbb{N}$.

Observation 1.10. This is nothing more than a general way to say that the Fourier transform of a smooth compactly supported function ψ decays super-polynomially. Indeed, taking $\lambda = 2\pi|\xi|$ and $\phi = \xi \cdot x/|\xi|$ we get

$$I(\lambda) = \hat{\psi}(\xi).$$

Proof. We consider the following vector field

$$L = \frac{1}{i\lambda}\sum_{k=1}^{d} a_k \frac{\partial}{\partial x_k} = \frac{1}{i\lambda}(a \cdot \nabla),$$

with $a = (a_1, \ldots, a_d) = \frac{\nabla\Phi}{|\nabla\Phi|^2}$. Then the transpose L^t of L is given by

$$L^t(f) = -\frac{1}{i\lambda}\sum_{k=1}^{d} \frac{\partial}{\partial x_k}(a_k f) = -\frac{1}{i\lambda}\nabla \cdot (af).$$

Because of our assumption on $\nabla\Phi$, the a_j and all their partial derivatives are each bounded on the support of ψ. Now observe that $L(e^{i\lambda\Phi}) = e^{i\lambda\Phi}$, therefore $L^N(e^{i\lambda\Phi}) = e^{i\lambda\Phi}$ for every positive integer N. Thus

$$I(\lambda) = \int_{\mathbb{R}^d} L^N(e^{i\lambda\Phi})\psi\,dx = \int_{\mathbb{R}^d} e^{i\lambda\Phi}(L^t)^N(\psi)\,dx.$$

Taking absolute values in the last integral gives $|I(\lambda)| \leq c_N\lambda^{-N}$ for positive λ, thus proving the lemma. □

Van der Corput's Lemma

More precise estimates can be obtained in the $d = 1$ case by evaluating

$$I(\lambda) = \int e^{i\lambda\Phi(x)}\, dx.$$

Lemma 1.11 (Van der Corput). *Suppose Φ is of class C^k in an interval $[a, b]$ with $k \geq 1$. Assume that $|\Phi^{(k)}(x)| \geq 1$ throughout the interval. Then*

$$\left| \int_a^b e^{i\lambda\Phi(x)}\, dx \right| \leq c_k \lambda^{-1/k}.$$

Let us analyze some special cases before proving the general result.

- **Φ is of class C^1**

Observation 1.12. Notice that a decay of order λ^{-1} is the best we can hope for, as follows by considering

$$\int_a^b e^{i\lambda x}\, dx = \frac{1}{i\lambda}(e^{i\lambda b} - e^{i\lambda a}).$$

Proof. Recall the operator L that occurred in the previous proposition. We may assume $\Phi' > 0$ on $[a, b]$, because when $\Phi' < 0$ we simply take complex conjugates. Similarly, we have

$$I(\lambda) = \int_a^b L(e^{i\lambda\Phi})\, dx = \int_a^b e^{i\lambda\Phi} L^t(1)\, dx + \left[e^{i\lambda\Phi} \frac{1}{i\lambda\Phi'} \right]_a^b.$$

Notice that, since the amplitude is not compactly supported, the boundary terms here play a role. $|\Phi'(x)| \geq 1$ so these two terms contribute a total bounded by $2/\lambda$. Moreover,

$$\int_a^b e^{i\lambda\Phi} L^t(1)\, dx \leq \int_a^b |L^t(1)|\, dx = \frac{1}{\lambda} \int_a^b \left| \frac{d}{dx}\left(\frac{1}{\Phi'}\right) \right| dx.$$

However, Φ' is monotonic, continuous and $|\Phi'(x)| \geq 1$, so $\frac{d}{dx}(1/\Phi')$ does not change sign in the interval $[a, b]$. Therefore,

$$\int_a^b \left| \frac{d}{dx}\left(\frac{1}{\Phi'}\right) \right| dx = \left| \int_a^b \frac{d}{dx}\left(\frac{1}{\Phi'}\right) dx \right| = \left| \frac{1}{\Phi'(b)} - \frac{1}{\Phi'(a)} \right|.$$

Altogether then $|I(\lambda)| \leq 3/\lambda$ as we wanted. $\square$

Remark 1.13. If we assume $|\Phi'(x)| \geq \mu$ (instead of $|\Phi'(x)| \geq 1$), then we get $|I(\lambda)| \leq c(\lambda\mu)^{-1}$. This reduces to replacing Φ by Φ/μ and λ by $\lambda\mu$.

- **Φ is of class C^2.**

Proof. We can assume that Φ'' is positive, hence $\Phi''(x) \geq 1 > 0$, otherwise complex conjugation will do the job. It follows that Φ' is strictly increasing and may have at most one critical point, which we call x_0. We split the domain of integration into three intervals:

$$I(\lambda) = \int_a^{x_0-\delta} e^{i\lambda\Phi(x)}\, dx + \int_{x_0-\delta}^{x_0+\delta} e^{i\lambda\Phi(x)}\, dx + \int_{x_0+\delta}^{b} e^{i\lambda\Phi(x)}\, dx =: I_1 + I_2 + I_3.$$

Since $\Phi'(x_0) = 0$, integrating the inequality $\Phi''(x) \geq 1$ between x_0 and $x_0 + \delta$, we deduce that $\Phi'(x) \geq \Phi'(x_0 + \delta) \geq \delta$ for all $x \in (x_0 + \delta, b)$. Thus, using what we have proven before, we obtain

$$|I_3(\lambda)| \leq \frac{c}{\delta\lambda}.$$

Similarly

$$|I_1(\lambda)| \leqslant \frac{c}{\delta\lambda}.$$

Moreover

$$|I_2(\lambda)| \leqslant \int_{x_0-\delta}^{x_0+\delta} dx = 2\delta.$$

Hence $|I(\lambda)| \leqslant 2\delta + 2c/(\delta\lambda)$. Minimizing in δ (taking $\delta^2 = c'/\lambda$) yields the proof. □

Now we can tackle the proof of the Van Der Corput's lemma for a general k.

Proof. We will prove the lemma by induction on k. The proof is a blueprint of the C^2 case. Assuming $\Phi^{(k)}(x) \geq 1$ we have that $\Phi^{(k-1)}$ is strictly increasing and has at most one critical point x_0. We split the domain of integration into three intervals:

$$I(\lambda) = \int_a^{x_0-\delta} e^{i\lambda\Phi(x)}\, dx + \int_{x_0-\delta}^{x_0+\delta} e^{i\lambda\Phi(x)}\, dx + \int_{x_0+\delta}^{b} e^{i\lambda\Phi(x)}\, dx =: I_1 + I_2 + I_3.$$

In both the first and third interval we have $|\Phi^{(k-1)}(x)| \geq \delta$, and the induction hypothesis gives

$$|I_1(\lambda)| \leq \frac{c_{k-1}}{(\delta\lambda)^{1/(k-1)}}, \qquad |I_3(\lambda)| \leq \frac{c_{k-1}}{(\delta\lambda)^{1/(k-1)}}.$$

As before, $|I_2(\lambda)| \leq 2\delta$. Combining and minimizing in δ yields the result. □

The Van der Corput's lemma extends naturally to the case where the amplitude is a function $\psi \in C^1(a, b)$.

Corollary 1.14. *Under the same hypotheses as in Lemma 1.11, if $\psi \in C^1(a, b)$, then*

$$\left| \int_a^b e^{i\lambda\Phi(x)}\psi(x)\, dx \right| \leq c_k \lambda^{-1/k}.$$

The proof exploits the previous lemma via integration by parts, we will not discuss the details but one can find them in [12].
In general, one may also get the complete asymptotic expansion of $I(\lambda)$.

Lemma 1.15. *Consider*
$$I(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda\Phi(x)}\psi(x)\,dx,$$
where ψ is a C^∞ function of compact support and $x = 0$ is the only critical point of Φ in the support of ψ, while $\Phi''(0) \neq 0$. Then for every positive integer N,
$$I(\lambda) = \frac{e^{i\lambda\Phi(0)}}{\lambda^{1/2}}\left(a_0 + a_1\lambda^{-1} + \cdots + a_N\lambda^{-N}\right) + O(\lambda^{-N-1/2}), \quad \text{as } \lambda \to \infty.$$
The a_k are determined by $\Phi''(0), \ldots, \Phi^{(2k+2)}(0)$, and $\psi(0), \ldots, \psi^{(2k)}(0)$. In particular
$$a_0 = \left(\frac{2\pi}{-i\Phi''(0)}\right)^{1/2}\psi(0)$$
.

Proof. The idea is to first treat the special case where $\Phi(x) = x^2$ using the properties of the Fourier transform. Indeed one can see that for $\Re(s) > 0$, the Fourier transform of the Gaussian gives
$$\int_{\mathbb{R}} e^{-\pi s x^2}\psi(x)\,dx = s^{-1/2}\int_{\mathbb{R}} e^{-\pi\xi^2/s}\widehat{\psi}(\xi)\,d\xi.$$
Both sides admit analytic continuation to the half-plane $\Re(s) > 0$. Setting
$$s = -\frac{i\lambda}{\pi},$$
we obtain
$$\int_{\mathbb{R}} e^{i\lambda x^2}\psi(x)\,dx = \left(\frac{\pi i}{\lambda}\right)^{1/2}\int_{\mathbb{R}} e^{-i\pi^2\xi^2/\lambda}\widehat{\psi}(\xi)\,d\xi.$$
Expanding the oscillatory factor we get
$$e^{-i\pi^2\xi^2/\lambda} = \sum_{k=0}^{N}\frac{(-i\pi^2\xi^2)^k}{k!\lambda^k} + O\left(\frac{|\xi|^{2N+2}}{\lambda^{N+1}}\right).$$
Substituting into the integral yields
$$\int_{\mathbb{R}} e^{i\lambda x^2}\psi(x)\,dx = \left(\frac{\pi i}{\lambda}\right)^{1/2}\sum_{k=0}^{N}\frac{(-i\pi^2)^k}{k!\lambda^k}\int_{\mathbb{R}}\xi^{2k}\widehat{\psi}(\xi)\,d\xi + O(\lambda^{-N-3/2}).$$
Using Fourier inversion we get
$$\int_{\mathbb{R}}\xi^{2k}\widehat{\psi}(\xi)\,d\xi = \frac{i^{2k}}{(2\pi)^{2k}}\psi^{(2k)}(0),$$
which gives the coefficients and completes the proof for $\Phi(x) = x^2$. The first term corresponds to $k = 0$ and is equal to
$$\left(\frac{\pi i}{\lambda}\right)^{1/2}\psi(0).$$
For a general Φ one requires much more calculations, but the proof really comes down to recalling the first case; Details can be found in [12] and [17]. $\square$

1.6 Fourier transform of the spherical measure

Finally we are able to perform all the computation needed to evaluate the Fourier transform of the spherical measure. Let σ be the surface measure on S^{d-1}. Its Fourier transform is defined by

$$\widehat{\sigma}(\xi) = \int_{S^{d-1}} e^{-2\pi i \xi\cdot\omega}\, d\sigma(\omega), \qquad \xi \in \mathbb{R}^d.$$

An important observation is that both the sphere and the surface measure are rotationally invariant. Indeed, if $R \in O(d)$ is an orthogonal transformation, then

$$\widehat{\sigma}(R\xi) = \int_{S^{d-1}} e^{-2\pi i R\xi\cdot\omega}\, d\sigma(\omega) = \int_{S^{d-1}} e^{-2\pi i \xi\cdot R^{-1}\omega}\, d\sigma(\omega) = \widehat{\sigma}(\xi),$$

since R preserves both the sphere and the measure $d\sigma$. Consequently, $\widehat{\sigma}(\xi)$ depends only on the Euclidean norm $|\xi|$. Then the direction of ξ can be chosen freely. So from now on

$$\xi = |\xi| e_d,$$

where $e_d = (0, \dots, 0, 1)$ is the last standard basis vector. This will allow us to reduce the problem of understanding the behavior of $\widehat{\sigma}(\xi)$ to estimating a one dimensional oscillatory integral.

Proposition 1.16. *For $|\xi| > 0$,*

$$\widehat{\sigma}(\xi) = \frac{2\pi}{|\xi|^{(d-2)/2}} J_{(d-2)/2}(2\pi|\xi|),$$

where J_ν denotes the Bessel function of the first kind of order ν.

Proof. Noting that $\xi = |\xi| e_d$ we get

$$\widehat{\sigma}(\xi) = \int_{S^{d-1}} e^{-2\pi i|\xi|\cos(\theta)} d\sigma(\omega),$$

where θ is the angle between the north pole e_d and w. Passing to spherical polar coordinates on S^{d-1} we can write

$$d\sigma = d\sigma_{d-1} = (\sin\theta)^{d-2} d\theta d\sigma_{d-2}.$$

Thus

$$\widehat{\sigma}(\xi) = |S^{d-2}| \int_0^\pi e^{-2\pi i|\xi|\cos\theta} (\sin\theta)^{d-2} d\theta.$$

Substituting $t = -\cos\theta$

$$\widehat{\sigma}(\xi) = |S^{d-2}| \int_{-1}^1 e^{-2\pi i|\xi|t} (1-t^2)^{(d-3)/2}\, dt.$$

The classical Poisson integral representation of the Bessel function states that for $\nu > -1/2$,

$$J_\nu(s) = \frac{(s/2)^\nu}{\sqrt{\pi}\,\Gamma\big(\nu+\frac{1}{2}\big)} \int_{-1}^1 e^{ist}(1-t^2)^{\nu-1/2}\, dt.$$

Applying this with

$$\nu = \frac{d-2}{2}, \qquad s = 2\pi|\xi|,$$

and simplifying the resulting constants using the duplication formula for the Gamma function, yields the claimed identity. $\square$

Bessel Functions

Besides the Poisson integral representation, Bessel functions also admit a Fourier-series representation. For $n \in \mathbb{Z}$ and $s \in \mathbb{R}$,

$$J_n(s) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{is\sin\theta} e^{-in\theta}\, d\theta.$$

To identify this representation with the Poisson integral formula, one observes that both families satisfy Bessel's differential equation

$$s^2 y'' + sy' + (s^2 - n^2)y = 0,$$

and the same recursion formulas

$$J_{n-1}(s) + J_{n+1}(s) = \frac{2n}{s} J_n(s), \qquad J_{n-1}(s) - J_{n+1}(s) = 2J_n'(s).$$

Together with the same initial normalization, this uniquely determines the functions. More precise details are given in [3]. To complete the estimates for $\widehat{\sigma}$ we will use the Fourier series representation.

Stationary Phase Analysis of the Bessel Function

We now apply the theory developed on oscillatory integrals to compute estimates for $J_n(s)$. Write

$$J_n(s) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{is\sin\theta} e^{-in\theta}\, d\theta.$$

The phase function is

$$\Phi(\theta) = \sin\theta,$$

whose critical points satisfy

$$\Phi'(\theta) = \cos\theta = 0.$$

Hence the two stationary points are

$$\theta_+ = \frac{\pi}{2}, \qquad \theta_- = -\frac{\pi}{2}.$$

Moreover,

$$\Phi''(\theta_+) = -1, \qquad \Phi''(\theta_-) = 1.$$

To justify the asymptotic expansion rigorously, one introduces a partition of unity adapted to neighborhoods of the two critical points:

$$1 = \chi_+ + \chi_- + \chi_0,$$

where $\chi_\pm$ are supported near $\theta_\pm$ and χ_0 avoids the critical set. The contribution of the non-stationary region associated with χ_0 is rapidly decaying due to Lemma 1.9. On the other hand, applying Lemma 1.11 and Lemma 1.15, near each critical point gives

$$\begin{aligned} J_n(s) &= \frac{1}{2\pi}\sqrt{\frac{2\pi}{s}}\left[e^{is}e^{-i\pi/4}e^{-in\pi/2} + e^{-is}e^{i\pi/4}e^{in\pi/2}\right] + O(s^{-3/2}) \\ &= \sqrt{\frac{2}{\pi s}}\cos\left(s - \frac{n\pi}{2} - \frac{\pi}{4}\right) + O(s^{-3/2}). \end{aligned}$$

We therefore obtain the classical asymptotic expansion.

Theorem 1.17. *For fixed $n \in \mathbb{Z}$ and $s \to +\infty$,*

$$J_n(s) = \sqrt{\frac{2}{\pi s}}\cos\left(s - \frac{n\pi}{2} - \frac{\pi}{4}\right) + O(s^{-3/2}).$$

Asymptotic Behaviour of the Spherical Fourier Transform

Combining Proposition 1.16 and the asymptotic formula just found, with

$$s = 2\pi|\xi|, \qquad n = \frac{d-2}{2},$$

we obtain

$$\widehat{\sigma}(\xi) = \frac{c_d}{|\xi|^{(d-1)/2}}\cos\left(2\pi|\xi| - \frac{(d-1)\pi}{4}\right) + O\left(|\xi|^{-(d+1)/2}\right).$$

In particular,

$$|\widehat{\sigma}(\xi)| = O\left(|\xi|^{-(d-1)/2}\right).$$

The stationary phase expansion, recall Lemma 1.15, also yields the fuller oscillatory asymptotic formula

$$\widehat{\sigma}(\xi) \sim |\xi|^{-(d-1)/2}\sum_{k=0}^{\infty}\frac{a_k^+ e^{2\pi i|\xi|} + a_k^- e^{-2\pi i|\xi|}}{|\xi|^k},$$

where the coefficients are determined by the higher derivatives arising in the local quadratic reductions near the two stationary points.

Chapter 2

The Spherical Maximal Theorem

In this chapter, we are going to introduce the spherical maximal function and prove the Spherical Maximal Theorem in dimension $d \geq 3$.

2.1 The Spherical Maximal Function

Let σ be the normalized surface measure on S^{d-1}. For $f \in L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, we define

$$A_t f(x) := \int_{S^{d-1}} f(x - ty)\, d\sigma(y). \tag{2.1}$$

Observation 2.1. Notice that we only assumed $f \in L^p(\mathbb{R}^d)$ but S^{d-1} is a set of Lebesgue measure 0 in $\mathbb{R}^d$, so one could change the value of $A_t f(x)$ arbitrarily and it would not even make sense to define spherical averages. For that reason we initially require $f \in C^\infty_{com}(\mathbb{R}^d)$. Density theorems will naturally extend our results to all $f \in L^p(\mathbb{R}^d)$; see [14].

We may write

$$A_t f(x) = f * \sigma_t(x),$$

where σ_t denotes the dilated measure, defined in the sense of distributions as:

$$\langle \sigma_t, f \rangle := \langle \sigma, f(\cdot\, t) \rangle.$$

As for the standard averaging operator defined on balls, we are interested in the conditions granting

$$\lim_{t \to 0} A_t f(x) = f(x).$$

It is then natural to introduce the maximal operator

$$Mf(x) := \sup_{t>0} |A_t f(x)|,$$

also called the *spherical maximal function*, and to understand for which $1 \leq p < \infty$ there exists a constant C_p such that

$$\|Mf\|_{L^p} \leq C_p \|f\|_{L^p} \quad \text{for all } f \in L^p(\mathbb{R}^d).$$

Note that fixed t the operator $f \to A_t f$ is bounded on $L^p(\mathbb{R}^d)$ for any $1 \leq p \leq \infty$. Since $A_t f$ is measurable, one can apply Minkowski's integral inequality

$$\|A_t f\|_{L^p} = \left(\int_{\mathbb{R}^d} \Big| \int_{S^{d-1}} f(x - ty) d\sigma(y) \Big|^p dx \right)^{1/p} \leq \int_{S^{d-1}} \left(\int_{R^d} |f(x - ty)|^p dx \right)^{1/p} d\sigma(y).$$

The invariance of the Lebesgue measure under translation then gives

$$\|A_t f\|_{L^p(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)}, \quad \text{for all } \ f \in L^p(\mathbb{R}^d).$$

It is easy to see that we cannot expect Mf to behave as well as the fixed-radius averages $A_t f$. For instance, let

$$f(x) := \chi_{B(0,1)}(x).$$

It follows that

$$Mf(x) \gtrsim \frac{1}{|x|^{d-1}} \quad \text{for any } x \in \mathbb{R}^d.$$

This implies

$$\int_{\mathbb{R}^d} |Mf|^p \gtrsim \int_{\mathbb{R}^d} \frac{1}{|x|^{p(d-1)}}.$$

But

$$\frac{1}{|x|^{p(d-1)}} \in L^1(\mathbb{R}^d) \quad \text{if and only if } \ p > \frac{d}{d-1}.$$

Thus, the best we can hope for is that M is bounded in the open range $p > \frac{d}{d-1}$. The two following theorem prove that this is exactly the case.

Theorem 2.2 (Stein, 1976)**.** *Let $d \geq 3$. For each $\frac{d}{d-1} < p \leq \infty$ there is a constant C_p such that*

$$\|Mf\|_{L^p(\mathbb{R}^d)} \leq C_p \|f\|_{L^p(\mathbb{R}^d)}$$

for all $f \in L^p(\mathbb{R}^d)$.

Stein's theorem, together with the preceding observations, gives a complete characterization of the L^p-boundedness of M. Note the restriction $d \geqslant 3$; it turns out that the case $d = 2$ is significantly more difficult to analyze. The L^p-boundedness of the circular maximal function was resolved by Bourgain ten years later.

Theorem 2.3 (Bourgain, 1986)**.** *Let $d = 2$, for all $2 < p \leqslant \infty$ there is a constant C_p such that*

$$\|Mf\|_{L^p(\mathbb{R}^2)} \leq C_p \|f\|_{L^p(\mathbb{R}^2)}$$

for all $f \in L^p(\mathbb{R}^2)$.

Historical note

Spherical averages often appear as solutions of partial differential equations. For instance, the spherical average

$$u(x,t) = \int_{\mathbf{S}^2} t\, f(x - ty)\, d\sigma(y)$$

is a solution of the following Cauchy problem for the *wave equation*

$$\begin{aligned} \Delta_x u(x,t) &= \frac{\partial^2 u}{\partial t^2}(x,t), \\ u(x,0) &= 0, \\ \frac{\partial u}{\partial t}(x,0) &= f(x), \end{aligned}$$

in $\mathbb{R}^3$. The introduction of the spherical maximal function is motivated by the fact that the related spherical average

$$u(x,t) = \int_{\mathbf{S}^2} f(x - ty)\, d\sigma(y)$$

solves the following Cauchy problem for*Darboux's equation*

$$\begin{aligned} \Delta_x u(x,t) &= \frac{\partial^2 u}{\partial t^2}(x,t) + \frac{2}{t}\frac{\partial u}{\partial t}(x,t), \\ u(x,0) &= f(x), \\ \frac{\partial u}{\partial t}(x,0) &= 0, \end{aligned}$$

in $\mathbb{R}^3$.

2.2 The Spherical Maximal Theorem for $d \geq 3$

In this section we focus on the complete proof of Stein's Theorem.

Theorem 2.4 (Stein, 1976). *Let $d \geq 3$. For each $\frac{d}{d-1} < p \leq \infty$ there is a constant C_p such that*

$$\|Mf\|_{L^p(\mathbb{R}^d)} \leq C_p \|f\|_{L^p(\mathbb{R}^d)}$$

for all $f \in L^p(\mathbb{R}^d)$.

Overview. The proof of the theorem provided by Stein in 1976 is of great interest due to its involvement of many classic techniques of harmonic analysis. In this chapter, we will attempt to unravel all the key steps of the original argument. The general strategy relies on the interpolation theorems stated in Chapter 1. By leveraging these results, one can reduce the proof to establishing two primary bounds:

- A strong L^2 estimate.
- A weak $(1,1)$ estimate.

The reduction to the L^2 space is particularly vital as it allows us to exploit the Fourier transform and its decay properties through Plancherel's theorem. Although the overall approach is consistent, there is a subtle but significant difference between the general case $d \geq 4$ and $d = 3$:

- For $d \geq 4$, a strong L^2 estimate can be obtained directly through an argument based on square functions.
- For $d = 3$, although the square function is still a winning heuristic, the proof requires a more nuanced approach. Specifically, one needs to first decompose the operator through a classical Littlewood-Paley decomposition.

2.3 Stein's proof of the L^2 bound for $d \geq 4$

We start by first analyzing the argument of the L^2 bound in dimension $d \geq 4$ to better understand the ideas of the complete proof

Observation 2.5. As noted before, it is far from obvious that $A_t f(x)$ is continuous in t for any $x \in \mathbb{R}^d$, or that

$$\sup_{t_1 < t < t_2} |A_t f(x)|$$

is even measurable in x. But we can avoid those problems by considering a priori estimates for well-behaved functions and extending them to all L^p by density arguments.

We introduce the square function

$$(Sf)(x) = \left(\int_0^\infty \left| \frac{\partial (A_t f)(x)}{\partial t} \right|^2 t \, dt \right)^{1/2},$$

which is certainly well-defined whenever $f \in C_c^1(\mathbb{R}^d)$. Denote by $\mathcal{M}$ the standard maximal operator, given in terms of averages over centered balls. The L^2 boundedness of the spherical maximal function in dimension $d \geq 4$ follows from these two propositions.

Proposition 2.6 (Pointwise estimate)**.** *For every $f \in C_c^1(\mathbb{R}^d)$ and every $x \in \mathbb{R}^d$,*

$$\sup_{t>0} |(A_t f)(x)| \leq (\mathcal{M} f)(x) + c\,(Sf)(x),$$

where $c = (2d)^{-1/2}$.

Proof. Fix $x \in \mathbb{R}^d$ and $t > 0$. Since $\lim_{s\to 0} s^d A_s f(x) = 0$, the fundamental theorem of calculus gives

$$\begin{aligned}
(A_t f)(x) &= t^{-d} \int_0^t \frac{\partial}{\partial s}[s^d (A_s f)(x)]\, ds \\
&= t^{-d} \cdot d \int_0^t s^{d-1} (A_s f)(x)\, ds + t^{-d} \int_0^t s^d \frac{\partial}{\partial s}[(A_s f)(x)]\, ds \\
&=: I_1 + I_2.
\end{aligned}$$

Estimate of I_1: Using spherical polar coordinates

$$\begin{aligned}
I_1 = t^{-d} \cdot d \int_0^t s^{d-1} (A_s f)(x)\, ds &= t^{-d} \cdot d \int_0^t s^{d-1} \int_{S^{d-1}} f(x - sy)\, d\sigma(y)\, ds \\
= t^{-d} \cdot \frac{d}{\omega_{d-1}} \int_{B(x,t)} f(y)\, dy \leq \mathcal{M} f(x),
\end{aligned}$$

Estimate of I_2: Applying the Cauchy-Schwarz inequality,

$$|I_2| \leq \left(\int_0^t s \left| \frac{\partial (A_s f)(x)}{\partial s} \right|^2 ds \right)^{1/2} \cdot t^{-d} \left(\int_0^t s^{2d-1}\, ds \right)^{1/2}.$$

Computing the second factor gives $c := t^{-d} \big(\int_0^t s^{2d-1}\, ds \big)^{1/2} = (2d)^{-1/2}$. Noting that the argument of the first integral is positive, we can let the upper limit of integration go to ∞, and get

$$|I_2| \leq cSf(x).$$

Combining both estimates and taking the supremum over $t > 0$ yields the proof □

Proposition 2.7 (L^2 bound for the square function)**.** *Let $d \geq 4$. Then there exists a constant $A > 0$ such that for every $f \in C_c^1(\mathbb{R}^d)$,*

$$\|Sf\|_{L^2(\mathbb{R}^d)} \leq A\, \|f\|_{L^2(\mathbb{R}^d)}.$$

Proof. Since $A_s f = f * \sigma_s$, taking the Fourier transform gives

$$\widehat{A_s f}(\xi) = \widehat{f}(\xi) \cdot m(s\xi), \quad \text{where} \quad m(\xi) = \widehat{\sigma}(\xi).$$

Differentiating in s

$$\left(\frac{\partial A_s f}{\partial s}\right)^{\wedge}(\xi) = \frac{\widetilde{m}(s\xi)}{s}\cdot \widehat{f}(\xi), \quad \text{where} \quad \widetilde{m}(\xi) = \sum_{j=1}^{d} \xi_j \frac{\partial m}{\partial \xi_j}.$$

Next, by Plancherel's theorem,

$$\int_{\mathbb{R}^d} \left| \frac{\partial (A_s f)(x)}{\partial s} \right|^2 dx = \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \frac{|\widetilde{m}(s\xi)|^2}{s^2} d\xi.$$

Tonelli's theorem gives

$$\|Sf\|_{L^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |(Sf)(x)|^2\, dx = \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \left(\int_0^\infty |\widetilde{m}(s\xi)|^2 \frac{ds}{s} \right) d\xi.$$

Therefore it suffices to show that the inner integral is uniformly bounded in ξ. From the estimates on $\widehat{\sigma}$ established in Chapter 1, we have

$$|\widetilde{m}(\xi)| \leq \min\left\{A'|\xi|,\ A'|\xi|^{-(d-3)/2}\right\}.$$

Splitting the integration at $s = |\xi|^{-1}$,

$$\int_0^\infty |\widetilde{m}(s\xi)|^2 \frac{ds}{s} \leq A' \left(|\xi|^2 \int_0^{|\xi|^{-1}} s\, ds + |\xi|^{-(d-3)} \int_{|\xi|^{-1}}^{\infty} s^{-(d-2)}\, ds \right).$$

Notice that the second integral on the right converges only when $d \geq 4$, and is then bounded by a constant multiple of $|\xi|^{d-3}$. Hence,

$$\int_0^\infty |\widetilde{m}(s\xi)|^2 \frac{ds}{s} \leq A,$$

uniformly in ξ, which yields the desired bound. □

Observation 2.8. The restriction $d \geq 4$ is sharp using this argument: for $d = 3$ the bound $|\widetilde{m}(\xi)| \leq A'|\xi|^{-(d-3)/2}$ provides no decay whatsoever, and the integral $\int_{|\xi|^{-1}}^{\infty} s^{-(d-2)}\, ds$ diverges.

Proof of the L^2 bound for M. Let $f \in C_c^1(\mathbb{R}^d)$. By Proposition 2.6,

$$\sup_{t>0} |A_t f(x)| \leq (\mathcal{M}f)(x) + c\,(Sf)(x).$$

Taking L^2 norms and using the triangle inequality,

$$\|Mf\|_{L^2} \leq \|\mathcal{M}f\|_{L^2} + c\,\|Sf\|_{L^2}.$$

Theorem 1.2 gives $\|\mathcal{M}f\|_{L^2} \leq C\|f\|_{L^2}$, and Proposition 2.7 gives $\|Sf\|_{L^2} \leq A\|f\|_{L^2}$. Combining,

$$\|Mf\|_{L^2} \lesssim \|f\|_{L^2},$$

which extends to all $f \in L^2(\mathbb{R}^d)$ by density. □

2.4 Proof of the L^p bound for $d \geq 3$

We now turn to the full proof of the L^p boundedness of the spherical maximal operator for the full range of exponents $p > d/(d-1)$. Since $A_t f = f * \sigma_t$ we can interpret Mf as a maximal Fourier multiplier operator defined as

$$\sup_{t>0} |(\widehat{f}(\xi)\, m(t\xi))^\vee|.$$

The strategy is to decompose the multiplier $m(\xi)$ into its radial components and estimate each piece separately.

2.4.1 Littlewood-Paley decomposition

We start by fixing a radial C^∞ function φ_0 in $\mathbb{R}^d$ such that $\varphi_0(\xi) = 1$ when $|\xi| \leq 1$ and $\varphi_0(\xi) = 0$ when $|\xi| \geq 2$. We define $\varphi(\xi) = \varphi_0(\xi) - \varphi_0(2\xi)$ and for $j \geq 1$ let

$$\varphi_j(\xi) = \varphi(2^{-j}\xi).$$

We observe that $\varphi_j(\xi)$ is localized near $|\xi| \approx 2^j$. Then we have

$$\sum_{j=0}^{\infty} \varphi_j = 1.$$

Set $m_j = \varphi_j m$ for all $j \geq 0$. The m_j's are C_0^∞ functions that satisfy $m = \sum_{j=0}^\infty m_j$. Setting

$$M_j f = \sup_{t>0} |(\widehat{f}(\xi)\, m_j(t\xi))^\vee|,$$

the triangle inequality gives the pointwise bound

$$M(f) \leq \sum_{j=0}^{\infty} M_j(f).$$

Our main goal from now on will be to estimate $M_j(f)$ for each j in a way that their L^p norms are summable.

Proposition 2.9 (Low-frequency)**.** *The operator M_0 is bounded on $L^p(\mathbb{R}^d)$ for all $1 < p \leq \infty$.*

Proof. Notice that

$$A_{0,t}(f)(x) = (\widehat{f}(\xi)\, m_0(t\xi))^\vee(x) = f * K_t^{(0)}(x), \quad \text{where} \quad K_t^{(0)}(x) = (m_0(t\xi))^\vee(x).$$

m_0 is a C^∞_{com} function, so its Fourier transform K^0 is a well-defined Schwartz function, hence it admits a radial integrable majorant $\mathcal{K}^{(0)}(x)$. From Lemma 1.6 it follows that:

$$\sup_{t>0} |(f * K_t^{(0)})(x)| \leq \mathcal{M}f(x) \cdot \int \mathcal{K}_t^{(0)}(y)\, dy \lesssim \mathcal{M}f(x).$$

The conclusion follows from the L^p boundedness of $\mathcal{M}$ $\square$

It remains to handle the pieces $M_j f$ for $j \geq 1$. Our goal will be to prove an estimate of the form:

$$\|M_j f\|_{L^p(\mathbb{R}^d)} \leq 2^{-j\varepsilon_d(p)} \|f\|_{L^p(\mathbb{R}^d)} \quad \text{where} \quad \varepsilon_d(p) > 0.$$

The idea is to follow a general principle of harmonic analysis, stating that it is best to interpolate before summing. We will prove:

- an L^2 bound $\|M_j f\|_{L^2} \leq C\, 2^{-j\alpha} \|f\|_{L^2}$ with a favorable constant $2^{-j\alpha}, \alpha > 0$;
- a weak $(1,1)$ bound $\|M_j f\|_{L^{1,\infty}} \leq C\, 2^j \|f\|_{L^1}$.

The Marcinkiewicz interpolation theorem will then give the L^p bound for $1 < p \leq 2$.

2.4.2 L^2 bound

We begin with an auxiliary proposition on square functions that will be used in the proof of the L^2 bound.

Proposition 2.10. *Suppose that the function m is supported in the annulus $R \leq |\xi| \leq 2R$ and is bounded. Then the square function*

$$G(f)(x) = \left(\int_0^\infty |(m(t\xi)\widehat{f}(\xi))^\vee(x)|^2 \frac{dt}{t} \right)^{\frac{1}{2}}$$

satisfies $\|Gf\|_{L^2(\mathbb{R}^d)} \lesssim \|m\|_{L^\infty} \|f\|_{L^2(\mathbb{R}^d)}$.

Proof. By Tonelli's theorem and Plancherel's theorem,

$$\|Gf\|_{L^2}^2 = \int_{\mathbb{R}^d} \int_0^\infty \left| (m(t\xi)\widehat{f}(\xi))^\vee(x) \right|^2 \frac{dt}{t}\, dx = \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \int_0^\infty |m(t\xi)|^2 \frac{dt}{t}\, d\xi.$$

Since m is supported in $R \leq |\xi| \leq 2R$, the inner integral is supported in $t \in [\frac{R}{|\xi|}, \frac{2R}{|\xi|}]$, giving

$$\int_0^\infty |m(t\xi)|^2 \frac{dt}{t} \leq \|m\|_{L^\infty}^2 \int_{R/|\xi|}^{2R/|\xi|} \frac{dt}{t} = \|m\|_{L^\infty}^2 \log 2.$$

Integrating against $|\widehat{f}(\xi)|^2$ and applying Plancherel's theorem yields the result. $\square$

Lemma 2.11. *There is a constant $C = C(d)$ such that for any $j \geq 1$,*

$$\|M_j(f)\|_{L^2(\mathbb{R}^d)} \leq C\, 2^{-j\frac{(d-2)}{2}} \|f\|_{L^2(\mathbb{R}^d)}$$

for all $f \in L^2(\mathbb{R}^d)$.

Proof. Define the square functions

$$G_j(f)(x) = \left(\int_0^\infty |A_{j,t}(f)(x)|^2 \frac{dt}{t} \right)^{\frac{1}{2}}, \quad \widetilde{G}_j(f)(x) = \left(\int_0^\infty |\widetilde{A}_{j,t}(f)(x)|^2 \frac{dt}{t} \right)^{\frac{1}{2}},$$

where

$$A_{j,t}f = (\widehat{f}(\xi)\, m_j(t\xi))^\vee \quad \text{and} \quad \widetilde{A}_{j,t}f = (\widehat{f}(\xi)\, \widetilde{m}_j(t\xi))^\vee$$

Recalling the definition of $\widetilde{m}(\xi) = \xi \cdot \nabla m(\xi)$, we have

$$s\frac{dA_{j,s}}{ds}(f) = \widetilde{A}_{j,s}(f)$$

Since $A_{j,s}(f)(x) = f * (m_j)_s^\vee(x)$ and

$$\int m_j^\vee(x)\, dx = \widehat{(m_j^\vee)}(0) = m_j(0) = 0,$$

by Lemma 1.6, it follows that

$$\lim_{s\to 0} A_{j,s}(f)(x) = 0 \quad \text{a.e.}$$

Step 1: Reduction to square functions: By the fundamental theorem of calculus

$$\begin{aligned}(A_{j,t}(f)(x))^2 &= \int_0^t \frac{d}{ds}(A_{j,s}(f)(x))^2\, ds \\ &= 2\int_0^t A_{j,s}(f)(x)\, \widetilde{A}_{j,s}(f)(x)\frac{ds}{s},\end{aligned}$$

from which we obtain the estimate

$$|A_{j,t}(f)(x)|^2 \leq 2\int_0^\infty |A_{j,s}(f)(x)|\, |\widetilde{A}_{j,s}(f)(x)|\frac{ds}{s}.$$

Taking the supremum over all $t > 0$ and integrating over $\mathbb{R}^d$, we obtain

$$\|M_j(f)\|_{L^2}^2 \leq 2\int_{\mathbb{R}^d}\int_0^\infty |A_{j,s}(f)(x)|\, |\widetilde{A}_{j,s}(f)(x)|\frac{ds}{s}\, dx.$$

Two applications of the Cauchy-Schwarz inequality (first in s, then in x) give

$$\begin{aligned}\|M_j(f)\|_{L^2}^2 &\leq 2\int_{\mathbb{R}^d} G_j(f)(x)\, \widetilde{G}_j(f)(x)\, dx \\ &\leq 2\|G_j(f)\|_{L^2}\|\widetilde{G}_j(f)\|_{L^2}. \end{aligned} \tag{2.2}$$

Step 2: Bounds on the square functions: Both m_j and $\widetilde{m}_j$ are supported in the annulus $2^{j-1} \leq |\xi| \leq 2^{j+1}$. Thus, by Proposition 2.10

$$\|G_j f\|_{L^2} \lesssim \|m_j\|_{L^\infty}\|f\|_{L^2}, \qquad \|\widetilde{G}_j f\|_{L^2} \lesssim \|\widetilde{m}_j\|_{L^\infty}\|f\|_{L^2}.$$

Moreover, the decay of $\widehat{\sigma}$ yields

$$\|m_j\|_{L^\infty} \lesssim 2^{-j(\frac{d-1}{2})}, \quad \|\widetilde{m}_j\|_{L^\infty} \lesssim 2^{j(1-\frac{d-1}{2})}.$$

Combining everything with (2.2) implies the desired bound:

$$\|M_j(f)\|_{L^2} \leq C\, 2^{-j\frac{(d-2)}{2}}\|f\|_{L^2}.$$

□

Remark 2.12. Notice that the series $\sum_{j=1}^\infty 2^{-j(d-2)/2}$ converges if and only if $d \geq 3$. This implies the L^2 boundedness of the spherical maximal operator for $d \geq 3$. Indeed, as we shall see, this property does not hold in dimension $d = 2$.

2.4.3 $L^{1,\infty}$ bound

Lemma 2.13. *There exists a constant $C = C(d) < \infty$ such that for all $j \geq 1$ and for all $f \in L^1(\mathbb{R}^d)$,*

$$\|M_j(f)\|_{L^{1,\infty}} \leq C\, 2^j \|f\|_{L^1}.$$

Proof. The proof of the lemma is based on the pointwise estimate

$$M_j(f) \leq C\, 2^j\, \mathcal{M}(f)$$

and the weak-type $(1,1)$ boundedness of the Hardy-Littlewood maximal operator $\mathcal{M}$. Let $K^{(j)} = (m_j)^\vee = \psi_{2^{-j}} * \sigma$, where $\psi := \varphi^\vee$ is a Schwartz function. Setting

$$(K^{(j)})_t(x) = t^{-d} K^{(j)}(t^{-1}x)$$

we have that

$$M_j(f) = \sup_{t>0} |(K^{(j)})_t * f|.$$

By Lemma 1.6 we just need to find an integrable radial majorant of $K^{(j)}$ with L^1 norm controlled by 2^j. We claim that for any $M > d$ there is a constant $C_M < \infty$ such that

$$|K^{(j)}(x)| = |(\psi_{2^{-j}} * \sigma)(x)| \leq \frac{C_M 2^j}{(1+|x|)^M}.$$

Since ψ is a Schwartz function, we have for every $N > 0$:

$$\|(1+|x|)^N \psi(x)\|_{L^\infty} \leq C_N,$$

which yields

$$|\psi_{2^{-j}} * \sigma(x)| \leq C_N \int_{S^{d-1}} \frac{2^{dj}\, d\sigma(y)}{(1+2^j|x-y|)^N}.$$

To bound the integral on the right-hand we split S^{d-1} into dyadic shells relative to x:

$$S_{-1}(x) = \mathbf{S}^{d-1} \cap \{y \in \mathbb{R}^d : 2^j|x-y| \leq 1\}$$

and for $k \geq 0$,

$$S_k(x) = \mathbf{S}^{d-1} \cap \{y \in \mathbb{R}^d : 2^k < 2^j|x-y| \leq 2^{k+1}\}.$$

The key observation is that for $y \in \mathbf{S}^{d-1}$,

$$\sigma(B(y,R) \cap \mathbf{S}^{d-1}) \lesssim R^{d-1}.$$

We get

$$\sigma(S_k(x)) \lesssim \begin{cases} 2^{(k+1-j)(d-1)} & k \leq j, \\ 1 & k > j, \end{cases}$$

Here, we chose the best bound for $k > j$. Accordingly, we split the integral as

$$\sum_{k=-1}^{j} \int_{S_k(x)} \frac{C_N 2^{dj}\, d\sigma(y)}{(1+2^j|x-y|)^N} + \sum_{k=j+1}^{\infty} \int_{S_k(x)} \frac{C_N 2^{dj}\, d\sigma(y)}{(1+2^j|x-y|)^N}.$$

First sum ($k \le j$)*:* the shells are contained in $B(0,3)$, so

$$\begin{aligned}\sum_{k=-1}^{j} \int_{S_k(x)} \frac{C_N 2^{dj}\, d\sigma(y)}{(1+2^j|x-y|)^N} &\lesssim 2^{dj} \sum_{k=-1}^{j} \frac{\sigma(S_k(x))\, \chi_{B(0,3)}(x)}{2^{kN}} \\ &\lesssim 2^{dj} \sum_{k=-1}^{j} 2^{(k+1-j)(d-1)} \cdot 2^{-kN}\, \chi_{B(0,3)}(x) \\ &\lesssim 2^j \chi_{B(0,3)}(x).\end{aligned}$$

Second sum ($k > j$)*:* using $\sigma(S_k(x)) \lesssim 1$ and the trivial bound $|x| \le 2^{k+2-j}$,

$$\begin{aligned}\sum_{k=j+1}^{\infty} \int_{S_k(x)} \frac{C_N 2^{dj}\, d\sigma(y)}{(1+2^j|x-y|)^N} &\lesssim 2^{dj} \sum_{k=j+1}^{\infty} \frac{\chi_{B(0,2^{k+2-j})}(x)}{2^{kN}} \\ &\lesssim 2^{dj} \sum_{k=j+1}^{\infty} 2^{-kN} \frac{(1+2^{k+2-j})^M}{(1+|x|)^M}.\end{aligned}$$

Combining the two estimates

$$\le C' \frac{2^j}{(1+|x|)^M} \left[1 + \sum_{k=j+1}^{\infty} \frac{2^{(k-j)(M-N)}}{2^{j(N+1-d)}} \right].$$

Since N can be chosen arbitrarily, we have $N > M > d$; the geometric series converges and gives

$$|(\psi_{2^{-j}} * \sigma)(x)| \le \frac{C_M\, 2^j}{(1+|x|)^M},$$

as claimed. □

2.4.4 Interpolation

We can now combine the two preceding lemmas via the Marcinkiewicz interpolation theorem to obtain the L^p bound for each piece M_j..

Proposition 2.14 (L^p bound for M_j)**.** *Let $1 < p \le 2$ and $d \ge 3$. There exists $C = C(d,p) < \infty$ such that for all $j \ge 1$,*

$$\|M_j f\|_{L^p(\mathbb{R}^d)} \le C\, 2^{-j\varepsilon_d(p)} \|f\|_{L^p(\mathbb{R}^d)},$$

where $\varepsilon_d(p) = (d-1) - \frac{d}{p}$.

Proof. We simply apply the Marcinkiewicz Interpolation theorem with $p_0 = 1$ and $p_1 = 2$. Lemma 2.11 gives the strong-type $(2,2)$ bound

$$\|M_j f\|_{L^2} \le C_2\, 2^{-j(d-2)/2} \|f\|_{L^2},$$

and Lemma 2.13 gives the weak-type $(1,1)$ bound

$$\|M_j f\|_{L^{1,\infty}} \leq C_1\, 2^j \|f\|_{L^1}.$$

Performing all the necessary calculations, we obtain $\varepsilon_d(p) = (d-1) - \frac{d}{p}$ which is positive if and only if:

$$p > \frac{d}{d-1},$$

which is precisely the threshold we were searching for. □

Putting everything together, we are finally able to prove Theorem 2.4.

Proof of Stein's Theorem. By the Littlewood–Paley decomposition

$$Mf \leq M_0 f + \sum_{j=1}^{\infty} M_j f.$$

By Proposition 2.9 and Proposition 2.14 we have

$$\|Mf\|_{L^p} \leq \|M_0 f\|_{L^p} + \sum_{j=1}^{\infty} \|M_j f\|_{L^p} \leq \left(C_0 + C' \sum_{j=1}^{\infty} 2^{-j\varepsilon_d(p)} \right) \|f\|_{L^p}.$$

Recall that $\varepsilon_d(p) > 0$ for $d/(d-1) < p \leq 2$. Thus, the series converges and we obtain:

$$\|Mf\|_{L^p} \leq C\|f\|_{L^p}, \quad \text{for } \frac{d}{d-1} < p \leq 2.$$

Since the operator M is trivially bounded on L^∞, to extend the theorem to $2 < p \leq \infty$, we interpolate between L^q and L^∞, for some $d/(d-1) < q < 2$. □

Chapter 3

A Geometric Approach

To gain a deeper understanding of the problem, we now adopt a more direct geometric approach. We will prove that the local version of the maximal operator is bounded on L^2 for all $d \geq 3$ using a purely geometric argument. The purpose of doing so is twofold. First, we will see a reduction of the L^p boundedness of the spherical maximal operator to a purely geometric inequality concerning sums of indicator functions of annuli. Second, it clarifies why the bound breaks for $d = 2$; highlighting the differences in tangency between $d \geq 3$ and $d = 2$. This chapter will be organized as follows.

- Reducing the problem to counting incidences between annuli
- Proving a spherical intersection bound.
- Establishing the complete proof of the L^2 boundedness of the "local" operator.
- Understanding what goes wrong in the plane.

3.1 Introducing a scale

Given $0 < \delta \ll 1$ we let $C^\delta(x, r)$ denote the annulus formed by the δ-neighborhood of $C(x, r)$, the sphere of radius r centered at $x \in \mathbb{R}^d$. For $f \in C(\mathbb{R}^d)$ we can extract spherical averages as limits of more "classical" averages:

$$Af(x, r) = \lim_{\delta \to 0^+} \frac{1}{|C^\delta(x, r)|} \int_{C^\delta(x,r)} f = \lim_{\delta \to 0^+} \fint_{C^\delta(x,r)} f.$$

We introduce the maximal operators

$$M^\delta f(x) := \sup_{r>0} \left| \fint_{C^\delta(x,r)} f \right|.$$

Proposition 3.1 (Reduction to M^δ)**.** *To prove the strong-type inequality*

$$\|Mf\|_{L^p(\mathbb{R}^d)} \leq C\|f\|_{L^p(\mathbb{R}^d)},$$

it suffices to show that the family of strong-type inequalities

$$\|M^\delta f\|_{L^p(\mathbb{R}^d)} \le C\|f\|_{L^p(\mathbb{R}^d)} \tag{3.1}$$

holds uniformly in $\delta \in (0,1)$.

Proof. Fix $x \in \mathbb{R}^d$, $r > 0$ and $\delta > 0$,

$$\left|\fint_{C^\delta(x,r)} f\right| \le \sup_{s>0}\left|\fint_{C^\delta(x,s)} f\right| = M^\delta f(x).$$

Taking the limit as $\delta \to 0^+$ we get $|Af(x,r)| \le \liminf_{\delta\to 0^+} M^\delta f(x)$; since this holds for every $r > 0$ we can take the supremum.

$$Mf(x) \le \liminf_{\delta\to 0^+} M^\delta f(x).$$

Now integrating the inequality and using Fatou's lemma together with the uniformity hypothesis yields the proof. $\square$

Remark 3.2. From now on we will simplify the setup by considering only the local maximal operators

$$Mf(x) := \sup_{1\le r\le 2} |Af(x,r)|, \qquad M^\delta f(x) := \sup_{1\le r\le 2}\left|\fint_{C^\delta(x,r)} f\right|.$$

The geometric arguments tend not to work well with the full range of radii; nevertheless, the Fourier analytic techniques developed later allow one to prove global maximal estimates as a consequence of local estimates. We will not go into more detail, but one can see here[15]

3.2 Discretization

Since $M^\delta f(x) \le M^\delta|f|(x)\ \forall x$ from now on we assume $f \ge 0$. A key feature of $M^\delta f$ is that it is essentially constant on scale δ, as made precise by the following lemma.

Lemma 3.3 (Locally constant property). *Let* $f \ge 0$. *If* $x, x' \in \mathbb{R}^d$ *with* $|x - x'| \le \delta$, *for* $\delta \in (0,1)$, *then*

$$M^\delta f(x) \lesssim M^\delta f(x').$$

Proof. Choose $r > 0$ such that

$$M^\delta f(x) \lesssim \left|\fint_{C^\delta(x,r)} f\right|.$$

Since $|x - x'| \le \delta$, we have the inclusion

$$C^\delta(x,r) \subseteq C^{2\delta}(x',r)$$

and furthermore we can write

$$C^{2\delta}(x', r) \subseteq C^{\delta}(x', r_1) \cup C^{\delta}(x', r_2),$$

where $r_1 = r - \delta$ and $r_2 = r + \delta$. Thus, since $f \geq 0$,

$$\begin{aligned} M^{\delta} f(x) &\lesssim \left| \fint_{C^{\delta}(x,r)} f \right| \lesssim \fint_{C^{2\delta}(x',r)} |f| \\ &\lesssim \fint_{C^{\delta}(x',r_1)} |f| + \fint_{C^{\delta}(x',r_2)} |f| \lesssim M^{\delta} f(x'). \end{aligned}$$

□

Remark 3.4. There are some technical issues regarding whether r_1, r_2 lie in the admissible range $[1, 2]$, but we will omit the details; see [5].

In view of Lemma 3.3, we can think of $M^{\delta} f$ as a discrete object: for all x belonging to some δ-cube Q centered at x_Q,

$$M^{\delta} f(x) \sim \left| \fint_{C^{\delta}(x_Q, r_Q)} f \right|.$$

Since the operator M^{δ} averages over a restricted range of radii, it is "local". This means that $M^{\delta} f(x)$ depends only on the values of f within a ball of finite radius centered at x. Consequently, to prove (3.1), it suffices to show

$$\|M^{\delta} f\|_{L^p(Q_0)} \lesssim \|f\|_{L^p(\mathbb{R}^d)},$$

whenever $Q_0 \subseteq \mathbb{R}^d$ is a cube of side-length 1. By translation-invariance, we can assume Q_0 is centered at 0.

Lemma 3.5 (Discretization)**.** *Let $\mathcal{Q}_{\delta}$ be a decomposition of Q_0 into subcubes Q_{δ} of side-length δ. For each Q_{δ} there exist $x_Q \in Q_{\delta}$ and $r_Q \in [1, 2]$ such that*

$$M^{\delta} f(x) \lesssim \left| \fint_{C^{\delta}(x_Q, r_Q)} f \right| \quad \textit{for all } x \in Q_{\delta}. \tag{3.2}$$

Consequently,

$$\|M^{\delta} f\|_{L^p(Q_0)} \lesssim \left(\sum_{C \in \mathcal{C}} \delta^d \left| \fint_{C^{\delta}} f \right|^p \right)^{1/p}, \tag{3.3}$$

where $\mathcal{C}$ is a collection of spheres with δ-separated centers lying in Q_0.

Proof. The first inequality is formal: since $f \in C(\mathbb{R}^d)$ implies $M^{\delta} f$ is bounded, and we can find $x_Q \in Q_{\delta}$ and $r_Q \in [1, 2]$ realizing a near-supremum. To prove (3.3), we write

$$\|M^{\delta} f\|_{L^p(Q_0)}^p = \sum_{Q_{\delta} \in \mathcal{Q}_{\delta}} \|M^{\delta} f\|_{L^p(Q_{\delta})}^p,$$

and apply (3.2) together with $|Q_\delta| \sim \delta^d$ to get

$$\|M^\delta f\|_{L^p(Q_0)} \lesssim \left(\sum_{C \in \mathcal{C}'} \delta^d \left| \fint_{C^\delta} f \right|^p \right)^{1/p} \tag{3.4}$$

where $\mathcal{C}'$ is a collection of spheres with centers lying in disjoint δ-cubes. To extract a subcollection $\mathcal{C} \subset \mathcal{C}'$ of spheres with δ-separated centers that satisfies (3.3), we employ a classic pigeonholing argument.Consider the grid of disjoint δ-cubes containing the centers of the spheres in $\mathcal{C}'$. We can naturally identify these cubes with a lattice of integer coordinates. Take 3^d colors and assign one to each cube according to the reduction modulo 3 of its coordinate tuple. Taking all cubes with the same color, we partition $\mathcal{C}'$ into 3^d disjoint subcollections with elements at distance at least 2δ. Splitting the total sum in (3.4) across these classes, we can write:

$$\sum_{C \in \mathcal{C}'} = \sum_{\text{color } 1} + \sum_{\text{color } 2} + \cdots + \sum_{\text{color } 3^d}$$

By the pigeonhole principle, there must exist at least one subcollection, which we will call $\mathcal{C}$, such that

$$\sum_{C \in \mathcal{C}} \delta^d \left| \fint_{C^\delta} f \right|^p \geq \frac{1}{3^d} \sum_{C \in \mathcal{C}'} \delta^d \left| \fint_{C^\delta} f \right|^p$$

Taking $\mathcal{C}$ as our desired subcollection concludes the proof.

□

3.3 Linearization

We exploit ℓ^p duality to linearize the right-hand side of (3.3).

Proposition 3.6 (Linearization)**.** *We have*

$$\|M^\delta f\|_{L^p(Q_0)} \lesssim \delta^{\frac{d}{p}-1} \Big\| \sum_{C \in \mathcal{C}} a_c \chi_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \|f\|_{L^p(\mathbb{R}^d)}, \tag{3.5}$$

for some sequence $(a_c)_{C \in \mathcal{C}}$ *satisfying* $\left(\sum_{C \in \mathcal{C}} |a_c|^{p'}\right)^{1/p'} \leq 1$.

Proof. Starting from (3.3), we write

$$\left(\sum_{C \in \mathcal{C}} \delta^d \left| \fint_{C^\delta} f \right|^p \right)^{1/p} \approx \delta^{\frac{d}{p}-1} \left(\sum_{C \in \mathcal{C}} \left| \int_{C^\delta} f \right|^p \right)^{1/p}.$$

Viewing this as an ℓ^p norm and applying ℓ^p duality, we obtain

$$= \delta^{\frac{d}{p}-1} \sum_{C \in \mathcal{C}} a_c \int_{C^\delta} f$$

for some $(a_c)_{C\in\mathcal{C}}$ with $\ell^{p'}$ norm at most 1. Interchanging summation and integration and applying Hölder's inequality,

$$\left|\sum_{C\in\mathcal{C}} a_c \int_{C^\delta} f\right| = \left|\int_{\mathbb{R}^d} \sum_{C\in\mathcal{C}} a_c \chi_{C^\delta} f\right| \leq \Big\| \sum_{C\in\mathcal{C}} a_c \chi_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \|f\|_{L^p(\mathbb{R}^d)},$$

which gives (3.5). $\square$

This reduces the proof of the L^p boundedness of the local spherical maximal operator to estimates such as

$$\Big\| \sum_{C\in\mathcal{C}} a_c \chi_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \lesssim \delta^{1-\frac{d}{p}} \big(\sum_{C\in\mathcal{C}} |a_c|^{p'}\big)^{1/p'} \tag{3.6}$$

holding for all $(a_c)_{C\in\mathcal{C}}$ and for any collection $\mathcal{C}$ of spheres with δ-separated centers lying in a cube.

Remark 3.7. Notice that (3.6) is a purely geometric inequality. Proving it for the general p would give a completely geometric proof of the Spherical Maximal Theorem, but as of today for $d \geq 3$ it is still an unsolved problem. However, it should be noted that, in a remarkable and intricate paper, Schlag [8] gave an entirely geometric proof of Bourgain's circular maximal theorem. In the following part, we will prove the estimate for $p = 2$ and $d \geq 3$, where the Hilbert structure of the space will make the proof much easier. This will be very instructive in developing the intuition that will be used in the next chapter.

3.4 Intersection bounds for annuli

Before proving the $p = 2$ case of (3.6) we will prove a volume bound for intersecting annuli.

Lemma 3.8 (Intersection lies in a strip)**.** *Let $x_1 \neq x_2 \in \mathbb{R}^d$ and $r_1, r_2 > 0$. If $y \in C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)$, then*

$$\left\langle y - \frac{x_1 + x_2}{2}, \frac{x_2 - x_1}{|x_2 - x_1|} \right\rangle = \frac{r_1^2 - r_2^2}{2|x_2 - x_1|} + O\left(\frac{\delta}{|x_2 - x_1|}\right).$$

In particular, $C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)$ lies within a strip S of thickness $O\big(\delta/|x_1 - x_2|\big)$ around the affine hyperplane, orthogonal to $x_2 - x_1$, determined by the equation above.

Proof. We first show that $C(x_1, r_1) \cap C(x_2, r_2)$ lies on the affine hyperplane normal to the line through the two centers. Let $y \in C(x_1, r_1) \cap C(x_2, r_2)$, so that

$$|y - x_1|^2 = r_1^2 \qquad \text{and} \qquad |y - x_2|^2 = r_2^2.$$

Expanding and subtracting:

$$2\langle y, x_2 - x_1\rangle + |x_1|^2 - |x_2|^2 = r_1^2 - r_2^2.$$

Using $|x_1|^2 - |x_2|^2 = \langle x_1 + x_2, x_1 - x_2 \rangle$ and rearranging yields

$$2\left\langle y - \frac{x_1 + x_2}{2},\ x_2 - x_1 \right\rangle = r_1^2 - r_2^2.$$

Dividing by $|x_2 - x_1| \neq 0$ gives the hyperplane equation for the exact spheres. For the fattened annuli, since $y \in C^\delta(x_i, r_i)$ we have $|y - x_i|^2 = r_i^2 + O(\delta)$, and the same calculation gives

$$2\left\langle y - \frac{x_1 + x_2}{2},\ x_2 - x_1 \right\rangle = r_1^2 - r_2^2 + O(\delta).$$

Normalizing by $|x_2 - x_1|$ gives the desired estimate. □

Lemma 3.9 (Sphere-slab area estimate)**.** *For $d \geq 3$, let $S = [a, a + \lambda] \times \mathbb{R}^{d-1}$ be a slab of width λ with $0 \leq a < a + \lambda \leq 1$. Then*

$$\mathrm{Area}^{d-1}\big(C(0, 1) \cap S\big) \lesssim \lambda.$$

Proof. We slice $C(0, 1)$ with the hyperplanes $H_t = \{t\} \times \mathbb{R}^{d-1}$. By the co-area formula:

$$\mathrm{Area}^{d-1}(C(0, 1) \cap S) = \int_a^{a+\lambda} \mathrm{Area}^{d-2}(C(0, 1) \cap H_t)\, \frac{dt}{(1 - t^2)^{1/2}}.$$

The slice $C(0, 1) \cap H_t$ is a $(d - 2)$-dimensional sphere of radius $(1 - t^2)^{1/2}$, so

$$\mathrm{Area}^{d-2}(C(0, 1) \cap H_t) \sim (1 - t^2)^{\frac{d-2}{2}}.$$

Substituting:

$$\mathrm{Area}^{d-1}(C(0, 1) \cap S) \sim \int_a^{a+\lambda} (1 - t^2)^{\frac{d-3}{2}}\, dt.$$

For $d \geq 3$ the exponent $(d - 3)/2 \geq 0$, so $(1 - t^2)^{(d-3)/2} \leq 1$ for all $t \in [-1, 1]$, giving

$$\mathrm{Area}^{d-1}(C(0, 1) \cap S) \lesssim \lambda.$$

□

Remark 3.10. Notice how for $d = 2$ the proof breaks down since:

$$(1 - t^2)^{\frac{d-3}{2}} = (1 - t^2)^{-\frac{1}{2}},$$

is no longer bounded on $[-1, 1]$.

Theorem 3.11 (Annulus intersection volume bound)**.** *For $d \geq 3$, $r_1, r_2 \in [1, 2]$*

$$\left|C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)\right| \lesssim \frac{\delta}{1 + \delta^{-1}|x_1 - x_2|}.$$

Proof. We split into two cases according to the separation of the centers.

Case 1: $|x_1 - x_2| \geq \delta$. We have

$$\big|C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)\big| \lesssim \delta \cdot \mathrm{Area}^{d-1}\big(C(x_1, r_1) \cap S\big),$$

where the factor δ accounts for the radial thickness of the annulus and S denotes the strip from Lemma 3.8, which by that same lemma has width $O(\delta/|x_1 - x_2|)$. Lemma 3.9 then gives

$$\big|C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)\big| \lesssim \delta \cdot \frac{\delta}{|x_1 - x_2|} = \frac{\delta^2}{|x_1 - x_2|} \sim \frac{\delta}{1 + \delta^{-1}|x_1 - x_2|},$$

where the last equivalence uses $|x_1 - x_2| \geq \delta$.

Case 2: $|x_1 - x_2| \leq \delta$.

The strip estimate degenerates, so we bound the intersection simply by the volume of a single fattened sphere:

$$\big|C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)\big| \leq \big|C^\delta(x_1, r_1)\big| \lesssim \delta.$$

Since $|x_1 - x_2| \leq \delta$ implies $1 + \delta^{-1}|x_1 - x_2| \lesssim 1$, the right-hand side satisfies $\delta/(1 + \delta^{-1}|x_1 - x_2|) \sim \delta$. Interpolating between the two cases yields the proof. □

3.5 The L^2 bound

We are now able to prove (3.6) for $p = 2$, thus proving the L^2 boundedness of the local spherical maximal operator.

Lemma 3.12. *For $d \geq 3$, let $\mathcal{C}$ be a collection of spheres in $\mathbb{R}^d$ with δ-separated centers lying in $[-\frac{1}{2}, \frac{1}{2}]^d$ and radii in $[1, 2]$; For any sequence of complex coefficients $(a_C)_{C \in \mathcal{C}}$, we have*

$$\Big\|\sum_{C \in \mathcal{C}} a_C \chi_{C^\delta}\Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1 - \frac{d}{2}} \left(\sum_{C \in \mathcal{C}} |a_C|^2\right)^{1/2}.$$

To avoid technicalities, we assume $a_C = 1$ for all $C \in \mathcal{C}$; the complete result requires only additional Cauchy–Schwarz estimates. Our goal is thus to show

$$\Big\|\sum_{C \in \mathcal{C}} \chi_{C^\delta}\Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1 - \frac{d}{2}} \# \mathcal{C}^{1/2}.$$

Proof. Since $L^2(\mathbb{R}^d)$ is a Hilbert space, we expand the norm as an inner product:

$$\begin{aligned}\Big\|\sum_{C \in \mathcal{C}} \chi_{C^\delta}\Big\|^2_{L^2(\mathbb{R}^d)} = \Big\langle \sum_{C_1 \in \mathcal{C}} \chi_{C_1^\delta}, \sum_{C_2 \in \mathcal{C}} \chi_{C_2^\delta}\Big\rangle &= \sum_{C_1, C_2 \in \mathcal{C}} \langle \chi_{C_1^\delta}, \chi_{C_2^\delta}\rangle \\ &= \sum_{C \in \mathcal{C}} \|\chi_{C^\delta}\|^2_{L^2(\mathbb{R}^d)} + \sum_{\substack{C_1, C_2 \in \mathcal{C} \\ C_1 \neq C_2}} \langle \chi_{C_1^\delta}, \chi_{C_2^\delta}\rangle\end{aligned}$$

Diagonal contribution. Since $|C^\delta| \sim \delta$,

$$\sum_{C \in \mathcal{C}} \|\chi_{C^\delta}\|_{L^2}^2 \sim \delta \, \#\mathcal{C},$$

which is already better than the required bound.

Off-diagonal contribution. Note that $\langle \chi_{C_1^\delta}, \chi_{C_2^\delta} \rangle = |C_1^\delta \cap C_2^\delta|$. Applying Theorem 3.11 and decomposing dyadically in the distance between centers:

$$\begin{aligned}
|\sum_{\substack{C_1, C_2 \in \mathcal{C} \\ C_1 \neq C_2}} \langle \chi_{C_1^\delta}, \chi_{C_2^\delta} \rangle| &\lesssim \delta^2 \sum_{\substack{C_1, C_2 \in \mathcal{C} \\ C_1 \neq C_2}} \mathrm{dist}(C_1, C_2)^{-1} \\
&\approx \delta^2 \sum_{\ell=0}^{\log_2 \delta^{-1}} \sum_{C_1 \in \mathcal{C}} \sum_{\substack{C_2 \in \mathcal{C} \\ \mathrm{dist}(C_1, C_2) \sim 2^\ell \delta}} \mathrm{dist}(C_1, C_2)^{-1} \\
&\lesssim \delta \sum_{\ell=0}^{\log_2 \delta^{-1}} 2^{-\ell} \sum_{C_1 \in \mathcal{C}} \#\{C_2 \in \mathcal{C} : \mathrm{dist}(C_1, C_2) \sim 2^\ell \delta\}.
\end{aligned}$$

We now use the counting estimate

$$\#\{C_2 \in \mathcal{C} : \mathrm{dist}(C_1, C_2) \sim 2^\ell \delta\} \lesssim 2^{\ell d} \quad \text{for all } C_1 \in \mathcal{C}, \tag{3.7}$$

This is a direct consequence of the δ-separation hypothesis satisfied by the elements of $\mathcal{C}$. Indeed, for a fixed C_1, the relevant C_2 centers must lie within an annulus of volume $\sim (2^\ell \delta)^d$; dividing this by the disjoint volume $\sim \delta^d$ occupied by each C_2 yields (3.7). It follows

$$\lesssim \delta \, \#\mathcal{C} \sum_{\ell=0}^{\log_2 \delta^{-1}} 2^{(d-1)\ell} \lesssim \delta^{-(d-2)} \, \#\mathcal{C}.$$

Combining both contributions,

$$\| \sum_{C \in \mathcal{C}} \chi_{C^\delta} \|_{L^2}^2 \lesssim \delta^{-(d-2)} \, \#\mathcal{C},$$

which, upon taking square roots, gives the lemma. □

3.6 What goes wrong for $d = 2$

If we examine the argument above in the case $d = 2$, we see that the proof breaks down in the sphere-slab estimate of Lemma 3.9, since the integrand becomes

$$(1 - t^2)^{\frac{d-3}{2}} = (1 - t^2)^{-\frac{1}{2}},$$

which is no longer bounded on $[-1, 1]$. It blows up when $t \to \pm 1$, corresponding to nearly tangent pairs of circles.

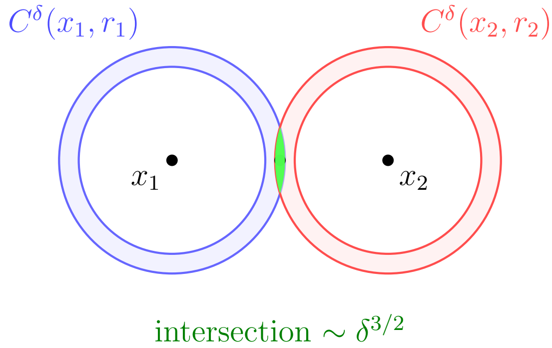

Figure 3.1: Two nearly tangent fattened circles in $\mathbb{R}^2$.

Near the tangency point, the arc over which the two circles are within distance δ of each other scales as $\delta^{1/2}$. Therefore, multiplying by the fattening thickness δ,

$$\left|C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)\right| \sim \delta^{1/2} \cdot \delta = \delta^{3/2}.$$

This is much larger than δ^2, which is the behavior expected from Theorem 3.11 considering $|x_1 - x_2|$. For $d \geq 3$, tangent intersections do not cause a problem. To see why, we compare the two regimes:

Tangent intersection ($|x_1 - x_2| \sim 1$, near-tangent): The intersection $C^\delta(x_1, r_1) \cap C^\delta(x_2, r_2)$ lies along a spherical cap of radius $\sim \delta^{1/2}$, i.e., a neighborhood of a *0-dimensional point* on S^{d-1} of radius $\sim \delta^{1/2}$ which has $d-1$ dimensional area $\sim \delta^{(d-1)/2}$. It follows that the volume of the fattened intersection is $\delta^{(d+1)/2}$.

Non-tangent intersection: The intersection is a δ-neighborhood of a $(d-2)$-dimensional sphere on S^{d-1}. Thus, the volume, determined purely by the co-dimension 2, is δ^2.

Although the intersection set is geometrically larger in the tangent case, it is of lower dimension (a 0-dimensional point versus a $(d-2)$-dimensional sphere). For $d \geq 3$, this dimensional reduction more than compensates.

Chapter 4

Combining Fourier Analysis and Geometry

In this chapter, combining the geometric intuition developed in Chapter 3 with the Fourier-analytic tools of Chapter 2 and 1, we give an alternative proof of Stein's theorem for the local version of the spherical maximal operator. This may seem redundant; however, the ideas that we are going to introduce are exactly the ones that, pushed further, allowed Bourgain to prove the theorem in the plane.

Overview. As in Chapter 2, we dyadically decompose the operator into frequency-localized pieces M_j, and we look for a bound of the form

$$\|M_j f\|_{L^p(\mathbb{R}^d)} \leq 2^{-j\varepsilon_d(p)} \|f\|_{L^p(\mathbb{R}^d)}. \tag{4.1}$$

However, this time, we do not pass through square functions. Instead, we exploit the "uncertainty principle". Since M_j is localized at frequency scale 2^j, it is almost constant at spatial scale 2^{-j}. This suggests to "break" M_j at scale $\delta := 2^{-j}$, replacing it with a sum of its elementary constituents $\tilde{\chi}_{C^\delta}$ indexed by spheres C of radius in $[1, 2]$ with δ-separated centers. Each $\tilde{\chi}_{C^\delta}$ is an L^∞-normalized function essentially supported on the δ-neighborhood C^δ of C, but – and this is the crucial difference compared to a purely geometric approach – it also carries some oscillation. Following the duality argument of Chapter 3, (4.1) is reduced to the bound

$$\Big\| \sum_{C \in \mathcal{C}} a_C \, \tilde{\chi}_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \lesssim \delta^{1-\frac{d}{p}+\varepsilon_d(p)} \Big(\sum_{C \in \mathcal{C}} |a_C|^{p'} \Big)^{1/p'}. \tag{4.2}$$

We obtain (4.2) by interpolating, via Riesz–Thorin, a trivial L^∞ bound against an L^2 bound carrying a substantial gain. The L^2 bound relies entirely on understanding the cross terms $\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle$ for two distinct spheres $C_1, C_2 \in \mathcal{C}$. Because $\tilde{\chi}_{C^\delta}$ oscillates, these cross terms are governed not just by how close C_1 and C_2 are, but by how *tangent* they are. This behavior is shown in the estimate.

Lemma 4.1 (Weak Orthogonality, informal statement)**.** *For every $N \in \mathbb{N}$,*

$$\big| \langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle \big| \lesssim_N \delta \big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d-1)/2} \big(1 + \delta^{-1} \Delta(C_1, C_2)\big)^{-N},$$

where $\Delta(C_1, C_2)$ quantifies the interior tangency of C_1 and C_2.

4.1 Discretization and Linearization

Recall from Chapter 2 that, by the Fourier inversion formula, the spherical means decompose in frequency as

$$Af(x,r) = \sum_{j=0}^{\infty} A_j f(x,r), \qquad A_j f(x,r) = \int e^{2\pi i x\cdot\xi}\, \hat{f}(\xi)\, \hat{\sigma}(r\xi)\, \varphi_j(\xi)\, d\xi,$$

where for $j \geq 1$ $\varphi_j(\xi) = \varphi(2^{-j}\xi)$. Equivalently, in physical space,

$$A_j f(x,r) = f * (\psi_j * \sigma_r)(x), \qquad \psi_0 = \varphi_0^{\vee}, \qquad \psi_j(x) = 2^{dj}\varphi^{\vee}(2^j x).$$

Let

$$M_j f(x) := \sup_{r\in[1,2]} |A_j f(x,r)|,$$

the frequency-localized local maximal operator that we wish to bound. The Low frequency contribution coming from $M_0 f$ are treated in the same way as in Chapter 2. Now let $\delta := 2^{-j}$ for $j \geq 1$, by the uncertainty principle, ψ_j, and hence the kernel $K_r^j := \psi_j * \sigma_r$, is locally constant at scale δ. This makes it natural to decompose space into cubes of side length δ. Let Q_0 be the cube of side-length $1/10$ centered at the origin, and let $\mathcal{Q}_j$ be the decomposition of Q_0 into dyadic cubes of side length δ.

Lemma 4.2 (Local constancy reduction)**.** *For each $Q \in \mathcal{Q}_j$ fix a center $x_Q \in Q$ and a radius $r_Q \in [1,2]$. Then*

$$\big\|M_j f\big\|_{L^p(Q_0)}^p \lesssim \sum_{Q\in\mathcal{Q}_j} \delta^d \Big| \int_{\mathbb{R}^d} f(y)\, \psi_j * \sigma_{r_Q}(x_Q - y)\, dy \Big|^p.$$

Proof. Since $\big\|M_j f\big\|_{L^p(Q_0)}^p = \sum_{Q\in\mathcal{Q}_j} \big\|M_j f\big\|_{L^p(Q)}^p$, it suffices to estimate each summand. On Q, local constancy of the kernel at scale δ shows that $A_j f(x,r)$ does not vary appreciably from $A_j f(x_Q, r_Q)$ as (x,r) ranges over $Q \times [1,2]$, so $\big\|M_j f\big\|_{L^p(Q)}^p \lesssim \delta^d\, |A_j f(x_Q, r_Q)|^p$. Summing over $Q \in \mathcal{Q}_j$ gives the claim. □

Definition 4.3 (Discrete pieces)**.** For a sphere $C = C(x,r)$ define

$$\tilde{\chi}_{C^\delta}(y) := \delta \cdot \psi_j * \sigma_r(x - y), \qquad \delta = 2^{-j}.$$

By the uncertainty principle $\tilde{\chi}_{C^\delta}$ should be thought of as supported on the δ-neighbourhood C^δ of C. What distinguishes $\tilde{\chi}_{C^\delta}$ from a mere indicator function is its oscillation: heuristically,

$$\tilde{\chi}_{C^\delta}(y) \approx e^{2\pi i \xi_C \cdot y} \chi_{C^\delta}(y).$$

This oscillation is the key feature that will let us upgrade a purely geometric estimate into the sharper weak orthogonality bound of Section §6. Arguing by duality as in Chapter 3, Lemma 4.2 reduces the proof of (4.1) to the following statement.

Lemma 4.4. *let $\mathcal{C}$ be a collection of spheres centered at δ-separated points of Q_0 with radii in $[1,2]$. For every $(a_C)_{C\in\mathcal{C}}$ we have*

$$\Big\| \sum_{C\in\mathcal{C}} a_C\, \tilde{\chi}_{C^\delta} \Big\|_{L^{p'}(\mathbb{R}^d)} \lesssim \delta^{1-d/p+\varepsilon_d(p)} \Big(\sum_{C\in\mathcal{C}} |a_C|^{p'} \Big)^{1/p'}. \tag{4.3}$$

This estimate is considerably easier to handle than its purely geometric counterpart (3.6) from Chapter 3. Moreover, the same arguments used to prove the L^2 bound adapt directly to this frequency-localized setting. In analogy with the classical proof of Stein the strategy is to first establish (4.3) at $p = 2$ with a strong choice of exponent $\varepsilon_d(2) > 0$, and then to interpolate this L^2 bound against a trivial L^∞ bound — leveraging the fact that $\varepsilon_d(2)$ is in fact larger than what is strictly needed to prove the L^2 estimate alone.

4.2 L^∞ Estimate

Lemma 4.5 (Frequency localized L^∞ bound)**.** *For $d \geq 2$ and $\mathcal{C}$ satisfying the hypothesis of Lemma 4.4,*

$$\Big\| \sum_{C\in\mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^\infty(\mathbb{R}^d)} \lesssim \delta^{1-d+\varepsilon_d(1)} \sup_{C\in\mathcal{C}} |a_C|, \qquad \varepsilon_d(1) := -1. \tag{4.4}$$

Proof. We bound

$$\Big\| \sum_{C\in\mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_\infty \leq \#\mathcal{C} \cdot \sup_{C\in\mathcal{C}} \big\| \tilde{\chi}_{C^\delta} \big\|_\infty \sup_{C\in\mathcal{C}} |a_C|.$$

The δ-separation of the centers gives $\#\mathcal{C} \lesssim \delta^{-d}$, so it suffices to show

$$\big\| \tilde{\chi}_{C^\delta} \big\|_\infty \lesssim 1 \quad \text{for all } C \in \mathcal{C}.$$

Recalling the definition of $\tilde{\chi}_{C^\delta}(x) := \delta \cdot (\psi_j * \sigma_r)(x)$ with $\psi_j(x) = \delta^{-d}\psi(\delta^{-1}x)$ for $\delta = 2^{-j}$ and C centered at the origin. We obtain

$$\tilde{\chi}_{C^\delta}(x) = \delta \cdot \int_{S^{d-1}} \delta^{-d} \psi\big(\delta^{-1}(x - r\theta)\big)\, d\sigma(\theta) = \delta^{1-d} \int_{S^{d-1}} \psi\big(\delta^{-1}(x - r\theta)\big)\, d\sigma(\theta).$$

Taking the absolute value and the supremum over $x \in \mathbb{R}^d$ and $1 \leq r \leq 2$, it follows that

$$\big\| \tilde{\chi}_{C^\delta} \big\|_\infty \leq \delta^{1-d} \sup_{x\in\mathbb{R}^d,\, 1\leq r\leq 2} \int_{S^{d-1}} \big|\psi\big(\delta^{-1}(x - r\theta)\big)\big|\, d\sigma(\theta).$$

Consequently, the uniform bound $\big\| \tilde{\chi}_{C^\delta} \big\|_\infty \lesssim 1$ reduces to proving the following shell estimate:

$$\sup_{x\in\mathbb{R}^d,\, 1\leq r\leq 2} \int_{S^{d-1}} |\psi(\delta^{-1}(x - r\theta))|\, d\sigma(\theta) \lesssim \delta^{d-1}. \tag{4.5}$$

To formally establish (4.5), we exploit the fact that ψ is the Fourier-Transform of a localized bump function, meaning that $\psi \in \mathcal{S}(\mathbb{R}^d)$. This implies that for any arbitrary integer $N \in \mathbb{N}$:

$$|\psi(y)| \leq \frac{C_N}{(1 + |y|)^N}, \qquad \forall y \in \mathbb{R}^d.$$

Setting $y = \delta^{-1}(x - r\theta)$, (4.5) comes down to

$$C_N \int_{S^{d-1}} \frac{1}{\big(1 + \delta^{-1}|x - r\theta|\big)^N} \, d\sigma(\theta) \lesssim \delta^{d-1}.$$

Substituting $\delta = 2^{-j}$, this bound is precisely the one obtained in Lemma 2.13 to prove the $L^{1,\infty}$ estimate. Hence it follows.

□

4.3 L^2 Theory

We now develop the L^2 theory of functions $\tilde{\chi}_{C^\delta}$, with the goal of proving (4.3) at $p = 2$ with a good choice of $\varepsilon_d(2)$. Expanding the squared norm gives

$$\Big\| \sum_{C \in \mathcal{C}} a_C \, \tilde{\chi}_{C^\delta} \Big\|^2_{L^2(\mathbb{R}^d)} = \sum_{C \in \mathcal{C}} |a_C|^2 \, \big\|\tilde{\chi}_{C^\delta}\big\|^2_{L^2(\mathbb{R}^d)} + \sum_{\substack{C_1, C_2 \in \mathcal{C} \\ C_1 \neq C_2}} a_{C_1} \, \overline{a_{C_2}} \, \langle \tilde{\chi}_{C_1^\delta}, \, \tilde{\chi}_{C_2^\delta} \rangle. \tag{4.6}$$

Diagonal Contribution

Lemma 4.6 (Diagonal estimate)**.** *Let $d \geq 2$ and $C \in \mathcal{C}$. Then*

$$\big\|\tilde{\chi}_{C^\delta}\big\|^2_{L^2(\mathbb{R}^d)} \lesssim \delta.$$

Proof. By Plancherel and the definition of $\tilde{\chi}_{C^\delta}$ and ψ_j,

$$\begin{aligned} \big\|\tilde{\chi}_{C^\delta}\big\|^2_{L^2(\mathbb{R}^d)} &= \delta^2 \int_{\mathbb{R}^d} |\hat{\sigma}_r(\xi)|^2 \, |\hat{\psi}_j(\xi)|^2 \, d\xi \\ &= \delta^2 \int_{\mathbb{R}^d} |\hat{\sigma}(r\xi)|^2 \, |\varphi(\delta|\xi|)|^2 \, d\xi. \end{aligned}$$

Recall from Chapter 1 that $\hat{\sigma}$ is radial and satisfies $|\hat{\sigma}(\xi)| \lesssim (1+|\xi|)^{-(d-1)/2}$. Since $\varphi(\delta|\xi|)$ is supported on $\delta^{-1}/2 \leq |\xi| \leq 2\delta^{-1}$ and $r \in [1, 2]$,

$$\begin{aligned} \big\|\tilde{\chi}_{C^\delta}\big\|^2_{L^2} &\lesssim \delta^2 \int_{\delta^{-1}/2 \leq |\xi| \leq 2\delta^{-1}} \big[(1 + |\xi|)^{-(d-1)/2}\big]^2 d\xi \\ &\sim \delta^2 \cdot \delta^{2(d-1)/2} \cdot \delta^{-d} = \delta, \end{aligned}$$

as required.

□

Off-Diagonal Contribution

We now estimate $|\langle \tilde{\chi}_{C_1^\delta}, \, \tilde{\chi}_{C_2^\delta} \rangle|$. Arguing with a slight variant of the proof above one readily obtains

$$|\langle \tilde{\chi}_{C_1^\delta}, \, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta\big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d-1)/2}, \tag{4.7}$$

where $\operatorname{dist}(C_1, C_2)$ is the distance between the centers of C_1 and C_2. This relies only on the decay of $\hat{\sigma}(\xi)$; however, it is not sufficient for our purposes. To gain a better bound we must also exploit the oscillation of $\hat{\sigma}$, which surprisingly involves the tangency between C_1 and C_2.

Definition 4.7 (Tangency between spheres)**.** For $C_\ell = C(x_\ell, r_\ell)$, $\ell = 1, 2$, define

$$\Delta(C_1, C_2) := \big||x_1 - x_2| - |r_1 - r_2|\big|.$$

The spheres C_1 and C_2 are interior tangent precisely when $\Delta(C_1, C_2) = 0$, i.e. $|x_1 - x_2| = |r_1 - r_2|$.

Lemma 4.8 (Refined off-diagonal estimate — Weak Orthogonality)**.** *Let $C_1, C_2 \in \mathcal{C}$ with $r_1, r_2 \in [1, 2]$ and $|x_1 - x_2| \lesssim 1$. For all $N \in \mathbb{N}$,*

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim_N \delta \big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d-1)/2} \big(1 + \delta^{-1} \Delta(C_1, C_2)\big)^{-N}. \tag{4.8}$$

In particular $\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle$ is negligible unless $\Delta(C_1, C_2) \lesssim \delta$.

Remark 4.9. If $r_1 = r_2$ then $\Delta(C_1, C_2) = \operatorname{dist}(C_1, C_2)$, and (4.8) yields

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim \delta \big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d+1)}, \tag{4.9}$$

which improves on (4.7). The same bound (4.9) continues to hold whenever $|r_1 - r_2| \lesssim \delta$.

We postpone the proof of Lemma 4.8 to Section §6, and first use it to establish the frequency-localized L^2 bound.

L^2 Estimate

Lemma 4.10 (Frequency localized L^2 estimate)**.** *Let $d \geq 2$ and let $\mathcal{C}$ be as in Lemma 4.4. Then*

$$\Big\| \sum_{C \in \mathcal{C}} a_C \, \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1-d/2+\varepsilon_d(2)} \Big(\sum_{C \in \mathcal{C}} |a_C|^2 \Big)^{1/2}, \qquad \varepsilon_d(2) := \frac{d-2}{2}. \tag{4.10}$$

To avoid technicalities we assume $a_C = 1$ for all $C \in \mathcal{C}$; the general case follows by an additional Cauchy–Schwarz argument. It suffices to prove

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \#\mathcal{C}^{1/2}. \tag{4.11}$$

Proof. We split $\sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta}$ according to the size of the radii, in order to exploit the strong decay in (4.9). Let $(I_k)_{k=1}^K$, $K \sim \delta^{-1}$, a partition of $[1, 2]$ into intervals of length $\sim \delta$, and for $1 \leq k \leq K$ set $\mathcal{C}_k := \{C \in \mathcal{C} : \operatorname{rad}(C) \in I_k\}$. By Cauchy–Schwarz,

$$\begin{aligned} \Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} &\leq \Big\| \sum_{k=1}^K \Big| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big| \Big\|_{L^2(\mathbb{R}^d)} \\ &\lesssim \delta^{-1/2} \Big(\sum_{k=1}^K \Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)}^2 \Big)^{1/2}. \end{aligned} \tag{4.12}$$

Fix k. Expanding the square as before,

$$\Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_2^2 = \sum_{C \in \mathcal{C}_k} \big\| \tilde{\chi}_{C^\delta} \big\|_2^2 + \sum_{\substack{C_1, C_2 \in \mathcal{C}_k \\ C_1 \neq C_2}} \langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle.$$

Diagonal terms are estimated by Lemma 4.6; off-diagonal terms are treated by combining (4.9) (valid since within $\mathcal{C}_k$ we have $|r_1 - r_2| \lesssim \delta$) with the dyadic counting argument used in the proof of Lemma 3.12. This gives

$$\begin{aligned} \Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)}^2 &\lesssim \delta \cdot \#\mathcal{C}_k + \delta \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} \sum_{C_1 \in \mathcal{C}_k} \sum_{\substack{C_2 \in \mathcal{C}_k \\ \operatorname{dist}(C_1, C_2) \sim 2^\ell \delta}} \big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d+1)} \\ &\lesssim \delta \cdot \#\mathcal{C}_k + \delta \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} \sum_{C_1 \in \mathcal{C}_k} \#\{C_2 \in \mathcal{C}_k : \operatorname{dist}(C_1, C_2) \sim 2^\ell \delta\} \, 2^{-(d+1)\ell} \\ &\lesssim \delta \#\mathcal{C}_k \Big(1 + \sum_{\ell=0}^{\lceil \log_2 \delta^{-1} \rceil} 2^{-\ell}\Big) \lesssim \delta \cdot \#\mathcal{C}_k, \end{aligned}$$

where in the second line we used

$$\#\{C_2 \in \mathcal{C} : \operatorname{dist}(C_1, C_2) \sim 2^\ell \delta\} \lesssim 2^{\ell d} \quad \text{for all } C_1 \in \mathcal{C}$$

Hence

$$\Big\| \sum_{C \in \mathcal{C}_k} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \delta^{1/2} [\#\mathcal{C}_k]^{1/2},$$

and combining this with (4.12),

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \Big(\sum_{k=1}^{K} \#\mathcal{C}_k \Big)^{1/2} = \#\mathcal{C}^{1/2},$$

which is (4.11). □

4.4 Interpolation and Stein's theorem

We now complete the proof of Lemma 4.4, by interpolating the L^∞ bound of Section §2 with the L^2 bound of Section of §3.

Lemma 4.11. *Let $d \geq 2$. For every collection $\mathcal{C}$ of spheres centered at δ-separated points of Q_0 with radii in $[1, 2]$, and every $(a_C)_{C \in \mathcal{C}}$, bound (4.3) holds for all $1 \leq p \leq 2$, with*

$$\varepsilon_d(p) := d - 1 - \frac{d}{p}.$$

Proof. By Lemma 4.5 (with $\varepsilon_d(1) = -1$),

$$\Big\| \sum_{C\in\mathcal{C}} a_C \tilde{\chi}_{C^\delta} \Big\|_{L^\infty(\mathbb{R}^d)} \lesssim \delta^{-d} \sup_{C\in\mathcal{C}} |a_C|,$$

and by Lemma 4.10 (with $\varepsilon_d(2) = \frac{d-2}{2}$),

$$\Big\| \sum_{C\in\mathcal{C}} a_C\, \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^d)} \lesssim \Big(\sum_{C\in\mathcal{C}} |a_C|^2 \Big)^{1/2}.$$

Both estimates hold for the same sequence $(a_C)_{C\in\mathcal{C}}$, so we may apply the Riesz–Thorin interpolation theorem to the linear map

$$T : (a_C)_{C\in\mathcal{C}} \longmapsto \sum_{C\in\mathcal{C}} a_C\, \tilde{\chi}_{C^\delta},$$

which is bounded from $\ell^\infty(\mathcal{C}) \to L^\infty(\mathbb{R}^d)$ with norm $\lesssim \delta^{-d}$, and from $\ell^2(\mathcal{C}) \to L^2(\mathbb{R}^d)$ with norm $\lesssim 1$. For $2 \le p' \le \infty$, write $\frac{1}{p'} = (1-\theta)\cdot 0 + \frac{\theta}{2} = \frac{\theta}{2}$ which implies $\frac{1}{p} = 1 - \frac{\theta}{2}$. Interpolation gives

$$\|T\|_{\ell^{p'} \to L^{p'}} \lesssim (\delta^{-d})^{1-\theta} \cdot 1^\theta = \delta^{-d(1-\theta)} = \delta^{-d(\frac{2}{p}-1)}.$$

Writing $-d\big(\frac{2}{p} - 1\big) = -\frac{2d}{p} + d$, and comparing with the claimed exponent $1 - \frac{d}{p} + \varepsilon_d(p)$, we recover

$$\varepsilon_d(p) = \Big(d - \frac{2d}{p}\Big) - 1 + \frac{d}{p} = d - 1 - \frac{d}{p},$$

as claimed. This proves Lemma 4.4. $\square$

Remark 4.12. For $d \ge 3$ we have $\varepsilon_d(p) > 0$ precisely when $p > \frac{d}{d-1}$, so Lemma 4.11 gives L^p-boundedness of the maximal spherical operator for $\frac{d}{d-1} < p \le 2$. The full range is recovered by interpolating M between L^q for $\frac{d}{d-1} < q < 2$ and the trivial bound $M : L^\infty \to L^\infty$. This completes the proof of Stein's theorem for the local maximal operator.

4.5 Weak Orthogonality: Proof of the Cross-terms Estimate

It remains to prove Lemma 4.8, restated here for convenience.

Lemma (Refined off-diagonal estimate — Weak Orthogonality)**.** *Let $C_1, C_2 \in \mathcal{C}$ with $r_1, r_2 \in [1,2]$, $|x_1 - x_2| \lesssim 1$ and $\delta = 2^{-j}$. For all $N \in \mathbb{N}$,*

$$|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle| \lesssim_N \delta \big(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\big)^{-(d-1)/2} \big(1 + \delta^{-1} \Delta(C_1, C_2)\big)^{-N}.$$

Proof. Recall from Chapter 1 that $\hat{\sigma}(\xi) = B_d(|\xi|)$ for a function $B_d : [0,\infty) \to \mathbb{C}$. For $C_\ell = C(x_\ell, r_\ell)$, $\ell = 1, 2$, we have $\tilde{\chi}_{C_\ell^\delta}(y) = \delta \cdot \psi_j * \sigma_{r_\ell}(x_\ell - y)$, so, taking the Fourier transform,

$$\begin{aligned}(\tilde{\chi}_{C_\ell^\delta})^\wedge(\xi) &= \delta \cdot e^{-2\pi i x_\ell \cdot \xi} \hat{\psi}_j(\xi) \hat{\sigma}_{r_\ell}(\xi) \\ &= \delta \cdot e^{-2\pi i x_\ell \cdot \xi} \varphi(\delta|\xi|) B_d(r_\ell|\xi|).\end{aligned}$$

By Plancherel's theorem and passing to polar coordinates,

$$\begin{aligned}\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle &= \delta^2 \int_{\mathbb{R}^d} B_d(r_1|\xi|) \overline{B_d(r_2|\xi|)} |\varphi(\delta|\xi|)|^2 e^{-2\pi i (x_1 - x_2)\cdot \xi} \, d\xi \\ &= \delta^2 \int_0^\infty B_d(r_1\rho)\, \overline{B_d(r_2\rho)}\, |\varphi(\delta\rho)|^2 \Big(\int_{S^{d-1}} e^{-2\pi i \rho (x_1 - x_2)\cdot\theta} \, d\sigma(\theta) \Big) \rho^{d-1} \, d\rho.\end{aligned}$$

Since

$$\int_{S^{d-1}} e^{-2\pi i \rho (x_1 - x_2)\cdot\theta} \, d\sigma(\theta) = \hat{\sigma}(\rho(x_1 - x_2)) = B_d(\rho|x_1 - x_2|),$$

we obtain

$$\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle = \delta^2 \int_0^\infty B_d(r_1\rho)\, \overline{B_d(r_2\rho)}\, B_d(\rho|x_1 - x_2|)\, |\varphi(\delta\rho)|^2 \, \rho^{d-1} \, d\rho. \tag{4.13}$$

Simply using the decay of B_d would only reproduce (4.7); to gain the extra factor in $\Delta(C_1, C_2)$ we must exploit the oscillation of B_d, recalling the stationary phase expansion from Chapter 1:

$$\hat{\sigma}(\xi) \sim |\xi|^{-(d-1)/2} \sum_{k=0}^\infty \frac{a_k^+ e^{2\pi i|\xi|} + a_k^- e^{-2\pi i|\xi|}}{|\xi|^k}.$$

This gives, for every M,

$$B_d(r) = \sum_{k=0}^{M} \Big(\sum_{\pm} a_{k,d}^\pm e^{\pm 2\pi i r} \Big) r^{-k-\frac{d-1}{2}} + r^{-M-\frac{d-1}{2}} E_{M,d}(r), \tag{4.14}$$

with

$$\left| \frac{\partial^k}{\partial r^k} E_{M,d}(r) \right| \lesssim_{M,d,k} 1 \qquad \text{for all } k \in \mathbb{N}_0 \text{ and } r \gtrsim 1.$$

We substitute (4.14) into (4.13), analyzing the resulting terms separately.

The leading term: Substituting

$$\begin{aligned}B_d(r_1\rho) &\longrightarrow a_{0,d}^+ e^{2\pi i r_1 \rho} (r_1\rho)^{-\frac{d-1}{2}}, \\ \overline{B_d(r_2\rho)} &\longrightarrow \overline{a_{0,d}^+}\, e^{-2\pi i r_2 \rho} (r_2\rho)^{-\frac{d-1}{2}}, \\ B_d(\rho|x_1 - x_2|) &\longrightarrow a_{0,d}^- e^{-2\pi i \rho|x_1 - x_2|} (\rho|x_1 - x_2|)^{-\frac{d-1}{2}},\end{aligned}$$

into (4.13) produces a term of the form

$$C(r_1,r_2)\cdot\delta^2\int_0^\infty e^{2\pi i(r_1-r_2-|x_1-x_2|)\rho}|\varphi(\delta\rho)|^2\rho^{-\frac{d-1}{2}}\,d\rho\cdot|x_1-x_2|^{-\frac{d-1}{2}} \tag{4.15}$$

$$=C(r_1,r_2)\cdot\delta\cdot\big(\delta^{-1}|x_1-x_2|\big)^{-\frac{d-1}{2}}\cdot(\mathcal{F}^{-1}\alpha)\big(\delta^{-1}(r_1-r_2-|x_1-x_2|)\big), \tag{4.16}$$

where $\alpha(\rho):=|\varphi(\rho)|^2\rho^{-\frac{d-1}{2}}$ is Schwartz (since φ vanishes near the origin), so $\mathcal{F}^{-1}\alpha$ is also Schwartz, and $C(r_1,r_2)$ is a constant depending only on r_1 and r_2. By the rapid decay of $\mathcal{F}^{-1}\alpha$, (4.16) is dominated, for every N, by

$$C_N\,\delta\,\big(\delta^{-1}|x_1-x_2|\big)^{-\frac{d-1}{2}}\big(\delta^{-1}\Delta(C_1,C_2)\big)^{-N}. \tag{4.17}$$

Other sign choices: The remaining contributions to the first-order term come from the other choices of sign in the phases

$$e^{\pm 2\pi i r_1\rho},\quad e^{\pm 2\pi i r_2\rho},\quad e^{\pm 2\pi i\rho|x_1-x_2|}.$$

Up to an overall sign (which reduces four of the eight combinations to the other four), all these choices give expressions identical to (4.16), with the argument of $\mathcal{F}^{-1}\alpha$ replaced by one of

$$\delta^{-1}(-r_1+r_2-|x_1-x_2|),\qquad \delta^{-1}(r_1+r_2+|x_1-x_2|),\qquad \delta^{-1}(r_1+r_2-|x_1-x_2|).$$

- For the first, writing $a=r_1-r_2$ and $b=|x_1-x_2|\geq 0$, the reverse triangle inequality gives

  $$|-r_1+r_2-|x_1-x_2||=|a+b|\geq\big||a|-b\big|=\big||r_1-r_2|-|x_1-x_2|\big|=\Delta(C_1,C_2),$$

  so this term is again bounded by (4.17).
- For the remaining two, using $r_1,r_2\in[1,2]$ and $|x_1-x_2|\lesssim 1$ (since C_1,C_2 have centers in Q_0),

  $$r_1+r_2+|x_1-x_2|\gtrsim 1,\qquad r_1+r_2-|x_1-x_2|\gtrsim 1,$$

  so, by the rapid decay of $\mathcal{F}^{-1}\alpha$, these contributions are $O(\delta^M)$ for every M and are therefore negligible.

Higher-order terms and the error. The same argument applies to every term $k\geq 1$ in expansion (4.14): each contributes a term dominated by (4.17) after replacing $\rho^{-\frac{d-1}{2}}$ with $\rho^{-\frac{d-1}{2}-k}$, which is still integrable against the Schwartz cutoff. Choosing M in (4.14) sufficiently large, the contribution of the error term $r^{-M-\frac{d-1}{2}}E_{M,d}$ is negligible as well.

Combining all contributions gives

$$|\langle\tilde{\chi}_{C_1^\delta},\tilde{\chi}_{C_2^\delta}\rangle|\lesssim_N\delta\cdot\big(\delta^{-1}\operatorname{dist}(C_1,C_2)\big)^{-\frac{d-1}{2}}\cdot\big(\delta^{-1}\Delta(C_1,C_2)\big)^{-N},$$

which, together with the trivial Cauchy–Schwarz bound $|\langle\tilde{\chi}_{C_1^\delta},\tilde{\chi}_{C_2^\delta}\rangle|\lesssim\delta$, yields

$$|\langle\tilde{\chi}_{C_1^\delta},\tilde{\chi}_{C_2^\delta}\rangle|\lesssim_N\delta\big(1+\delta^{-1}\operatorname{dist}(C_1,C_2)\big)^{-\frac{d-1}{2}}\cdot\big(1+\delta^{-1}\Delta(C_1,C_2)\big)^{-N}.$$

$\square$

4.6 Towards Bourgain's proof

This last section is dedicated to presenting some further ideas that explain how Bourgain extended the reductions carried out in this chapter to the plane, ultimately proving the Circular Maximal Theorem. We will not go into full detail; in particular, we stop short of a complete proof. Nevertheless, by the end of the section it will be clear what the last remaining obstacle is, and why overcoming it — via a careful analysis of the so-called "clam-shell" configuration — is the key step that Bourgain's original argument addresses.

We start by noting that both the L^2 and L^∞ estimates derived in §2 and §3 hold for $d \geq 2$. Furthermore, the exponent yielded by Lemma 4.11 is $\varepsilon_2(p) > 0$ for $p > 2$. However, this p falls strictly outside the admissible range $1 \leq p \leq 2$. This will be our starting point. By the previous results and restricted strong-type interpolation, it suffices to show the following:

Proposition 4.13. *For all $1 < q < 2$ there exists some exponent $\varepsilon(q) > 0$ such that*

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} \lesssim_q \delta^{2/q-1+\varepsilon(q)} \, \#\mathcal{C}^{1/q} \tag{4.18}$$

holds whenever $0 < \delta \ll 1$ and $\mathcal{C}$ is a collection of unit scale circles in the plane with δ-separated centers.

As before, a *"unit scale circle"* is a circle $C = C(x, r)$ in $\mathbb{R}^2$ with center $x \in [-1/10,\, 1/10]^2$ and radius $r \in [1, 2]$. From before,

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^2)} \lesssim \#\mathcal{C}^{1/2}, \tag{4.19}$$

and it is not difficult to show that

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^1(\mathbb{R}^2)} \leq \sum_{C \in \mathcal{C}} \|\tilde{\chi}_{C^\delta}\|_{L^1(\mathbb{R}^2)} \lesssim \delta \cdot \#\mathcal{C}. \tag{4.20}$$

Interpolating (4.19) and (4.20) via Hölder's inequality yields

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} \lesssim \delta^{2/q-1} \, \#\mathcal{C}^{1/q}, \qquad \text{for } 1 \leq q \leq 2,$$

which just misses (4.18) by an $\varepsilon(q)$-power of δ. Thus, our task is simply to obtain an ε-improvement over what we already have.

The first really important idea introduced by Bourgain is that one does not necessarily have to obtain estimates for the entire set $\mathcal{C}$. A better strategy is to partition $\mathcal{C}$ into subsets that behave slightly better in different Lebesgue spaces, and to interpolate on them beforehand.

Proposition 4.14. *Let $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$. If for a fixed $\varepsilon > 0$ we have*

$$\Big\| \sum_{C \in \mathcal{C}_2} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^2)} \lesssim \delta^{\varepsilon} \, \#\mathcal{C}^{1/2}, \tag{4.21}$$

and

$$\Big\| \sum_{C\in\mathcal{C}_1} \tilde{\chi}_{C^\delta} \Big\|_{L^1(\mathbb{R}^2)} \lesssim \delta^{1+\varepsilon}\, \#\mathcal{C}, \tag{4.22}$$

then Proposition 4.13 holds.

Proof. For any fixed $1 < q < 2$ we can apply the triangle inequality to bound

$$\Big\| \sum_{C\in\mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} \le \Big\| \sum_{C\in\mathcal{C}_1} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} + \Big\| \sum_{C\in\mathcal{C}_2} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)}.$$

Then we write

$$\frac{1}{q} = \frac{1-\theta}{1} + \frac{\theta}{2} \qquad \text{for some } 0 < \theta < 1.$$

Applying Hölder's inequality to interpolate between $L^1(\mathbb{R}^2)$ and $L^2(\mathbb{R}^2)$ gives

$$\begin{aligned}\Big\| \sum_{C\in\mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} \le{}& \Big\| \sum_{C\in\mathcal{C}_1} \tilde{\chi}_{C^\delta} \Big\|_{L^1(\mathbb{R}^2)}^{1-\theta} \Big\| \sum_{C\in\mathcal{C}_1} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^2)}^{\theta} \\ &+ \Big\| \sum_{C\in\mathcal{C}_2} \tilde{\chi}_{C^\delta} \Big\|_{L^1(\mathbb{R}^2)}^{1-\theta} \Big\| \sum_{C\in\mathcal{C}_2} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^2)}^{\theta}.\end{aligned}$$

Using (4.21), (4.22), (4.19) and (4.20) to estimate the contribution coming from different factors, one gets the following

$$\Big\| \sum_{C\in\mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^q(\mathbb{R}^2)} \lesssim \left(\delta^{(1+\varepsilon)(1-\theta)} + \delta^{(1-\theta)+\varepsilon\theta} \right) \#\mathcal{C}^{1/q}.$$

Taking

$$\varepsilon(q) := \varepsilon \cdot \min\left\{ \frac{2}{q} - 1,\ 2 - \frac{2}{q} \right\} > 0 \qquad \text{for } 1 < q < 2,$$

yields the proof.

□

Transverse Case

To further motivate the decomposition $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ we consider some special situations where it is possible to show a substantial gain in the exponent in a quite direct manner. Recall the weak-orthogonality estimate used to prove (4.19). We restate it again for convenience, assuming $d = 2$.

[**Weak orthogonality:**] For all pairs of unit scale circles C_1, C_2 and all $N \in \mathbb{N}$,

$$\big|\langle \tilde{\chi}_{C_1^\delta},\, \tilde{\chi}_{C_2^\delta} \rangle\big| \lesssim_N \delta \cdot \big(1 + \delta^{-1}\operatorname{dist}(C_1, C_2)\big)^{-1/2} \big(1 + \delta^{-1}\Delta(C_1, C_2)\big)^{-N} \tag{4.23}$$

Recall that $\Delta(C_1, C_2)$ measures the interior tangency between C_1 and C_2, so, given (4.23), we expect a significant gain in (4.19) if there are few pairs of tangent circles $(C_1, C_2) \in \mathcal{C} \times \mathcal{C}$. This is usually referred to as the "transverse case"; in the following lemma we make this statement rigorous.

Lemma 4.15 (Transverse case $\Rightarrow L^2$ improvement)**.** *Suppose $\mathcal{C}$ is a collection of unit scale circles with δ-separated centers which satisfies the additional hypothesis*

$$\#\{C_2 \in \mathcal{C} : \Delta(C_1, C_2) \le \delta^{1-\eta}\} \lesssim \delta^{2/3+4\varepsilon}\, \delta^{-2} \quad \textit{for all } C_1 \in \mathcal{C}, \tag{4.24}$$

for some fixed $0 < \varepsilon, \eta < 1$. Then

$$\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|_{L^2(\mathbb{R}^2)} \lesssim_{\eta,\varepsilon} \delta^{\varepsilon}\, [\#\mathcal{C}]^{1/2}.$$

Before giving a sketch of the proof, we will unravel the meaning of the hypothesis.

Remarks

- Recall that, by the uncertainty principle, since $\tilde{\chi}_{C^\delta}$ is frequency localized at scale $2^j = \delta^{-1}$, its support is essentially C^δ. Thus, δ is the smallest scale at which it "makes sense" to analyze the problem on the physical side. In particular, the condition
$$\Delta(C_1, C_2) \lesssim \delta$$
can be interpreted as "C_1^δ and C_2^δ are maximally tangent".
- Therefore the condition $\Delta(C_1, C_2) \le \delta^{1-\eta}$ means that C_1^δ and C_2^δ are tangent up to a relatively small tolerance factor of $\delta^{-\eta}$.
- Trivially,
$$\#\{C_2 \in \mathcal{C} : \Delta(C_1, C_2) \le \delta^{1-\eta}\} \le \#\mathcal{C} \le \delta^{-2}.$$
So hypothesis (4.24) represents a gain over the trivial bound by a factor of $\delta^{2/3+4\varepsilon}$. The exponent $\delta^{2/3}$ is not chosen at random: it is calibrated so that, after applying Hölder's inequality, it exactly compensates the pre-factor δ coming from (4.23).

Proof. Writing

$$\begin{aligned}\Big\| \sum_{C \in \mathcal{C}} \tilde{\chi}_{C^\delta} \Big\|^2_{L^2(\mathbb{R}^2)} &= \sum_{C_1 \in \mathcal{C}} \sum_{\substack{C_2 \in \mathcal{C} \\ \Delta(C_1,C_2) \le \delta^{1-\eta}}} \big\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \big\rangle + \sum_{C_1 \in \mathcal{C}} \sum_{\substack{C_2 \in \mathcal{C} \\ \Delta(C_1,C_2) > \delta^{1-\eta}}} \big\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \big\rangle \\ &=: \Sigma_{\text{tang}} + \Sigma_{\text{trans}}.\end{aligned}$$

We want to estimate the two contributions differently. For the transverse contribution, we will exploit the weak-orthogonality lemma. Since $\Delta(C_1, C_2) > \delta^{1-\eta}$, (4.23) implies

$$\big|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle\big| \lesssim_{N,\eta} \delta\big(1 + \delta^{-1}\Delta(C_1, C_2)\big)^{-\lceil N/\eta \rceil} \lesssim \delta^N \quad \text{for any } N \in \mathbb{N}.$$

Thus, Σ_{trans} is negligible. Instead, for the tangent contribution, our strategy is to use hypothesis (4.24) together with the weaker estimate

$$\big|\langle \tilde{\chi}_{C_1^\delta}, \tilde{\chi}_{C_2^\delta} \rangle\big| \lesssim \delta\big(1 + \delta^{-1}\,\mathrm{dist}(C_1, C_2)\big)^{-1/2}.$$

We first get

$$\Sigma_{\text{tang}} \lesssim \delta \sum_{C_1 \in \mathcal{C}} \sum_{\substack{C_2 \in \mathcal{C} \\ \Delta(C_1, C_2) \leq \delta^{1-\eta}}} \left(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\right)^{-1/2}.$$

Then, applying Hölder's inequality,

$$\Sigma_{\text{tang}} \lesssim \delta \sum_{C_1 \in \mathcal{C}} \#\{C_2 \in \mathcal{C} : \Delta(C_1, C_2) \leq \delta^{1-\eta}\}^{3/4} \left(\sum_{C_2 \in \mathcal{C}} \left(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\right)^{-2} \right)^{1/4}.$$

Decomposing dyadically on the distance set yields

$$\sum_{C_2 \in \mathcal{C}} \left(1 + \delta^{-1} \operatorname{dist}(C_1, C_2)\right)^{-2} \lesssim \log_2 \delta^{-1}.$$

Combining these observations leads us to

$$\Sigma_{\text{tang}} \lesssim \delta^{2\varepsilon} \, \#\mathcal{C},$$

which concludes the proof.

□

Tangent Case (Heuristic)

Having proved a stronger L^2 bound for a "transverse configuration" of circles, a natural question is what happens when hypothesis (4.24) fails. In particular, we consider a collection $\mathcal{C}$ such that there exists some $C_0 \in \mathcal{C}$ for which

$$\#\{C \in \mathcal{C} : \Delta(C_0, C) \leq \delta^{1-\eta}\} \quad \text{is large.}$$

All the circles $C \in \mathcal{C}$ are "almost tangent" to the fixed circle C_0. From now on we leave the rigorous setting and offer only a heuristic account, following [1]. What we are really interested in is the total number of "almost-tangent" pairs $(C_1, C_2) \in \mathcal{C} \times \mathcal{C}$. Trivially,

$$\#\{(C_1, C_2) \in \mathcal{C} \times \mathcal{C} : \Delta(C_1, C_2) \leq \delta^{1-\eta}\} \leq \#\mathcal{C}^2.$$

Furthermore, this bound can be attained if all the circles are arranged in a so-called "clam-shell" configuration, so that all the annuli are pairwise tangent along a common arc of C_0.

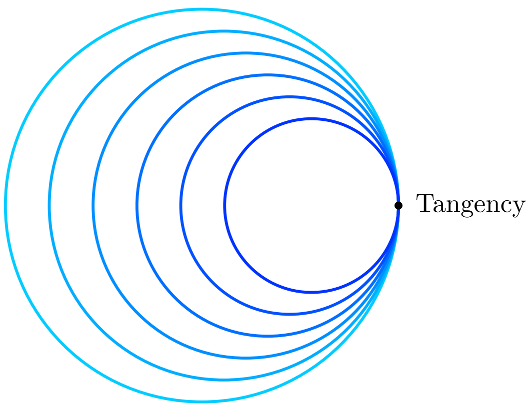

Figure 4.1: Clam-Shell Configuration

Surprisingly, it turns out that the "clam-shell" configuration is, in some sense, the only configuration in which there are many tangent pairs $(C_1, C_2) \in \mathcal{C} \times \mathcal{C}$. Proving a precise version of this statement is a key step of the full proof, and it actually requires only basic planar geometry; however, it is not in the scope of this work. What we have so far can be summarized as follows:

- By Lemma 4.15 we are led to consider the tangent case, in which all the circles in $\mathcal{C}$ are almost tangent to a fixed circle C_0.
- Again by Lemma 4.15, we can expect to be in a favorable situation unless $\mathcal{C} \times \mathcal{C}$ contains many pairs (C_1, C_2) of almost tangent circles.
- We are able to prove that the only way $\mathcal{C} \times \mathcal{C}$ can contain many pairs of almost tangent circles is if the circles in $\mathcal{C}$ are arranged in a clam-shell like configuration.

The only obstacle remaining is to characterize the "clam-shell" configuration precisely. What is needed is a stronger L^1 estimate — precisely the one obtained by Bourgain. Following in his footsteps would take us deep into the geometry of intersecting annuli, but we shall stop here. What is remarkable, though, is the simple beauty that emerges from such a complex problem, once the right reductions are made.

Bibliography

[1] Jean Bourgain. Averages in the plane over convex curves and maximal operators. *Journal d'Analyse Mathématique*, 47(1):69–85, 1986.

[2] Herbert Busemann and Willy Feller. Zur Differentiation der Lebesgueschen Integrale. *Fundamenta Mathematicae*, 22(1):226–256, 1934.

[3] Loukas Grafakos. *Classical Fourier Analysis*. Springer, 3rd edition, 2014.

[4] G. H. Hardy and J. E. Littlewood. A maximal theorem with function-theoretic applications. *Acta Mathematica*, 54:81–116, 1930.

[5] Jonathan Hickman. Personal webpage and academic resources. `https://webhomes.maths.ed.ac.uk/~jhickman/circ.html`, 2022. Accessed online.

[6] John M. Marstrand. Packing circles in the plane. *Proceedings of the London Mathematical Society*, s3-55(1):37–58, 1987.

[7] José L. Rubio de Francia. Maximal functions and Fourier transforms. *Duke Mathematical Journal*, 53(2):395–404, 1986.

[8] Wilhelm Schlag. A generalization of bourgain's circular maximal theorem. *Journal of the American Mathematical Society*, 10(1):103–122, 1997.

[9] Elias M. Stein. Maximal functions. i. spherical means. *Proceedings of the National Academy of Sciences*, 73(7):2174–2175, 1976.

[10] Elias M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

[11] Elias M. Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*. Princeton University Press, 2003.

[12] Elias M. Stein and Rami Shakarchi. *Functional Analysis: Introduction to Further Topics in Analysis*. Princeton University Press, 2011.

[13] Elias M. Stein and Jan-Olov Strömberg. Behavior of maximal functions in $\mathbb{R}^n$ for large n. *Arkiv för Matematik*, 21:259–269, 1983.

[14] Elias M. Stein and Stephen Wainger. Problems in harmonic analysis related to curvature. *Bulletin of the American Mathematical Society*, 84(6):1239–1295, 1978.

[15] Terence Tao. Lecture notes for Math 247B: Classical Fourier Analysis. https://www.math.ucla.edu/~tao/247b.1.07w/, 2019. University of California, Los Angeles (UCLA). Accessed online.

[16] Stephen Wainger. Some problems in harmonic analysis related to curves and surfaces. In Elias M. Stein, editor, *Beijing Lectures in Harmonic Analysis*, pages 41–75. Princeton University Press, 1986.

[17] Thomas H. Wolff. *Lectures on Harmonic Analysis.* American Mathematical Society, 2002. Edited by Izabella Łaba and Carol Shubin.