Harmonic analysis makes a fundamental use of divide-et-impera approaches. A particularly fruitful one is the decomposition of a function in terms of the frequencies that compose it, which is prominently incarnated in the theory of the Fourier transform and Fourier series. In many applications however it is not necessary or even useful to resolve the function \( f \) at the level of single frequencies and it suffices instead to consider how wildly different frequency components behave instead. One example of this is the (formal) decomposition of functions of \( \mathbb{R} \) given by

\[
f = \sum_{j \in \mathbb{Z}} \Delta_j f,
\]

where \( \Delta_j f \) denotes the operator

\[
\Delta_j f(x) := \int_{\{\xi \in \mathbb{R} : 2^j \leq |\xi| < 2^{j+1}\}} \hat{f}(\xi)e^{2\pi i \xi \cdot x} d\xi,
\]

commonly referred to as a (dyadic) frequency projection (you have encountered a similar one in the study of Hörmander-Mikhlin multipliers). Thus \( \Delta_j f \) represents the portion of \( f \) with frequencies of magnitude \( \sim 2^j \). The Fourier inversion formula can be used to justify the above decomposition if, for example, \( f \in L^2(\mathbb{R}) \).

Heuristically, since any two \( \Delta_j f, \Delta_k f \) oscillate at significantly different frequencies when \( |j - k| \) is large, we would expect that for most \( x \)’s the different contributions to the sum cancel out more or less randomly; a probabilistic argument typical of random walks (see Exercise 1) leads to the conjecture that \( |f| \) should behave “most of the time” like \( \left( \sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \right)^{1/2} \) (the last expression is an example of a square function). While this is not true in a pointwise sense, we will see in these notes that the two are indeed interchangeable from the point of view of \( L^p \)-norms: more precisely, we will show that for any \( 1 < p < \infty \) it holds that

\[
\|f\|_{L^p(\mathbb{R})} \sim_p \left( \sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \right)^{1/2} \left\| \left( \sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R})}.
\]

This is a result historically due to Littlewood and Paley, which explains the name given to the related theory. It is easy to see that the \( p = 2 \) case is obvious thanks to Plancherel’s theorem, to which the statement is essentially equivalent. Therefore one could interpret the above as a substitute for Plancherel’s theorem in generic \( L^p \) spaces when \( p \neq 2 \).

In developing a framework that allows to prove (†) we will encounter some variants of the square function above, including ones with smoother frequency projections that are useful in a variety of contexts. We will moreover show some applications of the above fact and its variants. One of these applications will be a proof of the boundedness of the spherical maximal function \( M_{S^{d-1}} \) introduced in a previous lecture.

**Notation:** We will use \( A \lesssim B \) to denote the estimate \( A \leq CB \) where \( C > 0 \) is some absolute constant, and \( A \sim B \) to denote the fact that \( A \lesssim B \lesssim A \). If the constant \( C \) depends on a list of parameters \( L \) we will write \( A \lesssim_L B \).
1. Motivation: estimates for the heat equation with sources

In this section we present certain natural objects that cannot be dealt with by (euclidean) Calderón-Zygmund theory alone, as a way to motivate the study of Littlewood-Paley theory.

Consider the heat equation on \( \mathbb{R}^d \)
\[
\begin{aligned}
&\partial_t u - \Delta u = f, \\
&u(x, 0) = f_0(x),
\end{aligned}
\] (1.1)
with \( f = f(x, t) \) representing the instantaneous heat generated by the source at point \( x \) at instant \( t \). If the total heat introduced by the sources is small, we expect to see that the solution \( u \) evolves slowly (that is, \( \partial_t u \) is also small). One sense in which this conjecture can be made rigorous is to control the \( L^p(\mathbb{R}^d \times \mathbb{R}) \)-norms of \( \partial_t u \) by the corresponding norms of the data \( f \). If we take a Fourier transform in both \( x \) and \( t \) we have
\[
\begin{aligned}
\hat{\partial_t u}(\xi, \tau) &= 2\pi i\tau \hat{u}(\xi, \tau), \\
\hat{\Delta u}(\xi, \tau) &= -4\pi^2|\xi|^2 \hat{u}(\xi, \tau).
\end{aligned}
\]

Now, we are trying to relate \( \partial_t u \) and \( f \) on the Fourier side (that is, we are trying to use spectral calculus). The simplest way to do so is to observe that, by taking the Fourier transform of both sides of (1.1), we have
\[\hat{\partial_t u} - \hat{\Delta u} = \hat{f},\]
that is
\[ (2\pi i\tau + 4\pi^2|\xi|^2) \hat{u}(\xi, \tau) = \hat{f}(\xi, \tau); \]
thus we can write with a little algebra
\[
\begin{aligned}
\partial_t \hat{u}(\xi, \tau) &= 2\pi i\tau \hat{u} = 2\pi i\tau \frac{2\pi i\tau + 4\pi^2|\xi|^2}{2\pi i\tau + 4\pi^2|\xi|^2} \hat{u} = \frac{\tau}{\tau - 2\pi i|\xi|^2} \hat{f}.
\end{aligned}
\]
Let \( m(\xi, \tau) := \frac{\tau}{\tau - 2\pi i|\xi|^2} \), so that the result can be restated as \( \partial_t \hat{u} = m \hat{f} \). Observe that both the real and imaginary parts of \( m \) are discontinuous at \( (0, 0) \) and nowhere else. By Fourier inversion we can interpret this equality as defining a linear operator \( T_m \), given by
\[ T_m g(x, t) = \int_{\mathbb{R}^d \times \mathbb{R}} m(\xi, \tau) \hat{g}(\xi, \tau) e^{i\xi \cdot x + i\tau \cdot t} d\xi d\tau; \]
thus we have come to realise that the partial derivative \( \partial_t u \) can be obtained (at least formally) from the function \( f \) by \( \partial_t u = T_m f \). To bound \( \|\partial_t u\|_{L^p(\mathbb{R}^d \times \mathbb{R})} \) is therefore equivalent to bounding \( \|T_m f\|_{L^p(\mathbb{R}^d \times \mathbb{R})} \) if \( T_m \) is bounded from \( L^p \) into \( L^p \) then we will have
\[ \|\partial_t u\|_{L^p(\mathbb{R}^d \times \mathbb{R})} \leq \|T_m f\|_{L^p(\mathbb{R}^d \times \mathbb{R})} \lesssim \|f\|_{L^p(\mathbb{R}^d \times \mathbb{R})}, \]
which is the type of control we are looking for.

Observe that by the properties of the Fourier transform our operator is actually a convolution operator: precisely\(^1\)
\[ T_m f(x, t) = f * \hat{m}(x, t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} f(x - y, t - s) \hat{m}(y, s) dy ds. \]

If we knew that the convolution kernel \( \hat{m} \) satisfied \( \|\hat{m}\|_{L^1} < \infty \) we would be done, since we could appeal to Young’s inequality. Unfortunately, this is not the case. We have already encountered this issue in dealing with singular integrals, so the natural thing to do is to check whether \( \hat{m} \) is a Calderón-Zygmund kernel. However, this

\(^1\) With a slight abuse of notation we have used \( \hat{m} \) to denote the inverse Fourier transform as well, in place of the more usual \( m \), motivated by the fact that on \( \mathbb{R}^d \) the two differ only by a reflection. We will continue to do so in the rest of these notes as it will always be clear from context which one is meant.
is also not the case! Indeed, if \( K \) is a (euclidean) Calderón-Zygmund convolution kernel, you have seen that it must satisfy properties such as \(|\nabla K(x)| \lesssim |x|^{-d-1}|\). However, \( \hat{m} \) enjoys a certain (parabolic) scaling invariance that makes it incompatible with the last inequality: indeed, one can easily see that for any \( \lambda > 0 \)

\[
\hat{m}(y, s) = \lambda^{d+2} \hat{m}(\lambda y, \lambda^2 s)
\]

and deduce from this a similar anisotropic rescaling invariance for \( \nabla \hat{m} \). Since \( |(y, s)|^{-d-1} \) is not invariant with respect to this rescaling, the inequality \( |\nabla \hat{m}(y, s)| \lesssim |(y, s)|^{-d-1} \) cannot possibly hold for all \( y, s \).

One solution to this issue is to develop a theory of parabolic singular integrals that extends the results of classical Calderón-Zygmund theory to those kernels that satisfy anisotropic rescaling invariance identities of the kind above. While this is a viable approach, we will take a different one in these notes. The approach we will take will not make use of said identities and will thus be more general in nature. With the methods of next section we will be able to show that \( T_m \) is indeed \( L^p \rightarrow L^p \) bounded for all \( 1 < p < \infty \).

2. Littlewood-Paley Theory

We begin by considering the frequency projections introduced in the beginning. Let \( I \) be an interval and define the frequency projection \( \Delta_I \) to be

\[
\Delta_I f(x) := \int 1_I(\xi)\hat{f}(\xi)e^{2\pi i \xi x}d\xi.
\]

An equivalent way of defining this is to work directly on the frequency side of things and stipulate that \( \Delta_I \) is the operator that satisfies

\[
\Delta_I \hat{f}(\xi) = 1_I(\xi)\hat{f}(\xi);
\]

this is well-defined for \( f \in L^2 \) and can be therefore extended to functions \( f \in L^2 \cap L^p \). We claim that \( \Delta_I \) defines an \( L^p \rightarrow L^p \) bounded operator for any \( 1 < p < \infty \) and any (possibly semi-infinite) interval \( I \) and, importantly, that its \( L^p \rightarrow L^p \) norm is bounded independently of \( I \). The point is that the operator \( \Delta_I \) is essentially a linear combination of two (modulated) Hilbert transforms, that are bounded in the same stated range. Indeed, observe that the Fourier transform of the Hilbert transform kernel \( \text{p.v.} 1/x \) is \(-i\pi \text{sign} (\xi)\); it is easy to show that, if \( I = (a, b) \), then

\[
1_{(a, b)}(\xi) = \frac{\text{sign}(\xi - a) - \text{sign}(\xi - b)}{2}.
\]

Now, the Fourier transform exchanges translations with modulations, and specifically \( \hat{f}(\xi + \theta) = (f(\cdot)e^{-2\pi i \theta \cdot})\hat{\chi}(\xi) \); we see therefore that we can write, with \( H \) the Hilbert transform,

\[
-i\pi \int \hat{f}(\xi)\text{sign}(\xi - a)e^{2\pi i \xi x} d\xi = -i\pi e^{2\pi i ax} \int \hat{f}(\xi + a)\text{sign}(\xi)e^{2\pi i \xi x} d\xi = e^{2\pi i ax} H(e^{-2\pi i \theta \cdot} f(\cdot))(x).
\]

If we let \( \text{Mod}_\theta \) denote the modulation operator \( \text{Mod}_\theta f(x) := e^{-2\pi i \theta x} f(x) \), we have therefore shown that

\[
-i\pi \Delta_{(a,b)} = \frac{1}{2} (\text{Mod}_{-a} \circ H \circ \text{Mod}_a - \text{Mod}_{-b} \circ H \circ \text{Mod}_b),
\]

and from the \( L^p \rightarrow L^p \) boundedness of \( H \) it follows that

\[
\|\Delta_I f\|_{L^p} \lesssim_p \|f\|_{L^p}
\]

(since \( \text{Mod}_\theta \) does not change the \( L^p \) norms at all).
At this point, another important thing to notice is that when \( I, J \) are disjoint intervals, then the operators \( \Delta_I, \Delta_J \) are orthogonal to each other. Indeed, one has by Parseval’s identity
\[
\langle \Delta_I f, \Delta_J g \rangle = \langle \Delta_I \hat{f}, \Delta_J \hat{g} \rangle = \int 1_I(\xi)\hat{f}(\xi)1_J(\xi)\overline{\hat{g}(\xi)}d\xi = \int 1_{I\cap J}(\xi)\hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi = 0.
\]
A consequence of this fact, together with Plancherel’s theorem, is the \( p = 2 \) case of the result that we mentioned in the introduction, that is inequality (1) and actually, one has more, namely that for any collection \( I \) of disjoint intervals that partition \( \mathbb{R} \) one has
\[
\left\| \left( \sum_{I \in \mathcal{I}} |\Delta_I f|^2 \right)^{1/2} \right\|_{L^2(\mathbb{R})} \leq \|f\|_{L^2}.
\]
Prove this in Exercise 2 to familiarise yourself with frequency projections.

2.1. Vector-valued estimates. As a step towards proving (1) (and also important for the applications), it will be useful to study vector-valued analogues of these inequalities, which are easier to prove in general. We change the setting to \( \mathbb{R}^d \) to work in more generality: so, if \( R \) is an axis-parallel rectangle (that is, of the form \( I_1 \times \ldots \times I_d \) with \( I_j \) intervals, possibly semi-infinite) in \( \mathbb{R}^d \), we define the frequency projection \( \Delta_R \) to be the operator that satisfies for any \( f \in L^2(\mathbb{R}^d) \)
\[
\Delta_R f(\xi) = 1_R(\xi)\hat{f}(\xi).
\]
The following is an abstract vector-valued analogue of (2.1).

**Proposition 2.1** (Vector-valued square functions). Let \( (\Gamma, d\mu) \) be a measure space with \( d\mu \) a \( \sigma \)-finite measure, and let \( \mathcal{H} \) denote the Hilbert space \( L^2(\Gamma, d\mu) \), that is the space of functions \( G : \Gamma \to \mathbb{C} \) such that \( \|G\|_{\mathcal{H}} := (\int |G(\gamma)|^2 d\mu(\gamma))^{1/2} < \infty \). We let \( (R_{\gamma})_{\gamma \in \Gamma} \) be a measurable collection of arbitrary rectangles of \( \mathbb{R}^d \) (that is, the mapping \( \Gamma : \gamma \to R_{\gamma} \) is a measurable function).

Given any vector-valued function \( f(x) = (f_{\gamma}(x))_{\gamma \in \Gamma} \in L^2(\mathbb{R}^d; \mathcal{H}) \) (that is, a function of \( \mathbb{R}^d \) that takes values in \( \mathcal{H} \)), we can define its vector-valued frequency projection \( \Delta_{\mathcal{R}} \) by
\[
\Delta_{\mathcal{R}} f := (\Delta_{R_{\gamma}} f_{\gamma})_{\gamma \in \Gamma}.
\]
Then we have that for any \( 1 < p < \infty \) it holds for all functions \( f = (f_{\gamma})_{\gamma \in \Gamma} \in L^p(\mathbb{R}^d; \mathcal{H}) \) that
\[
\|\Delta_{\mathcal{R}} f\|_{L^p(\mathbb{R}^d; \mathcal{H})} \lesssim_p \|f\|_{L^p(\mathbb{R}^d; \mathcal{H})}.
\]
In particular, the constant is independent of the collection of rectangles and even of the measure space \( (\Gamma, d\mu) \).

The level of abstraction in the previous statement is quite high, and it might take the reader a while to fully unpack its meaning. The following special case of Proposition 2.1 should help:

**Corollary 2.2.** Let \( (R_j)_{j \in \mathbb{N}} \) be a collection of arbitrary axis-parallel rectangles in \( \mathbb{R}^d \). If \( f = (f_j)_{j \in \mathbb{N}} \) is a sequence of functions such that \( (\sum_j |f_j|^2)^{1/2} \) is in \( L^p(\mathbb{R}^d) \), we have
\[
\left\| \left( \sum_j |\Delta_{R_j} f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^d)} \lesssim_p \left\| \left( \sum_j |f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^d)},
\]

\(^2\)Notice the space of rectangles of \( \mathbb{R}^d \) can be identified with \( \mathbb{R}^d \times \mathbb{R}^d \).
with constant independent of the collection of rectangles.

To prove (1) we will only need Corollary 2.2, but for the application of Littlewood-Paley theory we will give in Section 3.1 we will need the full power of Proposition 2.7. Prove in Exercise 3 that the corollary is indeed just a special case of the proposition, before looking at the proof below.

Remark 1. It is important to notice that there is no assumption on the family of rectangles other than the fact that they are all axis-parallel - in particular, they are not necessarily pairwise disjoint or even distinct.

Proof of Proposition 2.7 The proof is a vector-valued generalisation of what we said for frequency projections over intervals at the beginning of this section. First consider $d = 1$ and introduce the vector-valued Hilbert transform $H$ given by

$$
Hf := (Hf_{\gamma})_{\gamma \in \Gamma};
$$

we claim that this operator is $L^{p}(\mathbb{R}; \mathcal{M}) \to L^{p}(\mathbb{R}; \mathcal{N})$ bounded. Indeed, $H$ is given by convolution with the vector-valued kernel $K(x) := (p.v.1/x)_{\gamma \in \Gamma}$, and this kernel satisfies (as you can easily verify)

i) $\|K(x)\|_{L^{p}(\mathcal{M}; \mathcal{N})} \lesssim |x|^{-1}$,

ii) $\int_{|x|>2|y|} \|K(x-y) - K(x)\|_{L^{p}(\mathcal{M}; \mathcal{N})}dx \lesssim 1$,

iii) $\int_{|x|<|y|<R} K(x)dx = 0$ for any $0 < r < R$.

Then i)-ii)-iii) imply $L^{p}(\mathbb{R}; \mathcal{M})$ boundedness simply by vector-valued Calderón-Zygmund theory.

Next, observe that if $d = 1$ and the intervals $I = (I_{\gamma})_{\gamma \in \Gamma}$ are arbitrary, we can use the same trick used to prove (2.1) (rewriting $\Delta_{I_{\gamma}}$ as sum of Hilbert transforms conjugated with a modulation) to deduce boundedness of the corresponding vector-valued operator $\Delta_{\gamma}$ from the boundedness of $H$ above.

Finally, in the case of general $d > 1$ one should observe that if $R = I_{1} \times \ldots \times I_{d}$ then $\Delta_{R}$ factorises as

$$
\Delta_{R} = \Delta_{I_{1}}^{(1)} \circ \ldots \circ \Delta_{I_{d}}^{(d)},
$$

where $\Delta_{I_{\gamma}}^{(k)}$ denotes frequency projection in the $x_{k}$ variable. Therefore, writing $R_{R} = I_{1} \times \ldots \times I_{d}$ and $x = (x_{1}, x') \in \mathbb{R} \times \mathbb{R}^{d-1}$, we have by the one-dimensional result applied to $x_{1}$ that

$$
\|\Delta_{R}f\|_{L^{p}(\mathbb{R}^{d}; \mathcal{M})}^{p} = \iint \left\| (\Delta_{R_{R}}f_{\gamma}(x_{1}, x'))_{\gamma \in \Gamma} \right\|_{\mathcal{M}}^{p} dx_{1}dx'.
$$

iterating the argument for each variable, we obtain the claimed boundedness for generic $d$.

2.2. Smooth square function. In this subsection we will consider a variant of the square function appearing at the right-hand side of (1) where we replace the frequency projections $\Delta_{\gamma}$ by better behaved ones.

Let $\psi$ denote a smooth function with the properties that $\psi$ is compactly supported in the intervals $[-4, -1/2] \cup [1/2, 4]$ and is identically equal to 1 on the

\[\]
intervals $[-2, -1] \cup [1, 2]$. We define the smooth frequency projections $\widetilde{\Delta}_j$ by stipulating

$$\widetilde{\Delta}_j f(\xi) := \psi(2^{-j} \xi) \hat{f}(\xi);$$

notice that the function $\psi(2^{-j} \xi)$ is supported in $[-2^{j+2}, -2^{j-1}] \cup [2^{j-1}, 2^{j+2}]$ and identically 1 in $[-2^{j+1}, -2^j] \cup [2^j, 2^{j+1}]$. The reason why such projections are better behaved resides in the fact that the functions $\psi(2^{-j} \xi)$ are now smooth, unlike the characteristic functions $1_{[2^j, 2^{j+1}]}$. Indeed, they are actually Schwartz functions and you can see by Fourier inversion formula that $\Delta_j f = f * (2^j \hat{\psi}(2^j \cdot))$; the convolution kernel $2^j \hat{\psi}(2^j \cdot)$ is uniformly in $L^1$ and therefore the operator is trivially $L^p \to L^p$ bounded for any $1 \leq p \leq \infty$ by Young’s inequality, without having to resort to the boundedness of the Hilbert transform.

We will show that the following smooth analogue of (one half of) (1) is true (you can study the other half in Exercise [9]).

**Proposition 2.3.** Let $\widetilde{S}$ denote the square function

$$\widetilde{S} f := \Big( \sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \Big)^{1/2}.$$

Then for any $1 < p < \infty$ we have that the inequality

$$\|\widetilde{S} f\|_{L^p(\mathbb{R})} \lesssim_p \|f\|_{L^p(\mathbb{R})}$$

holds for any $f \in L^p(\mathbb{R})$.

We will give two proofs of this fact, to illustrate different techniques. We remark that the boundedness will depend on the smoothness and the support properties of $\psi$ only, and as such extends to a larger class of square functions.

**First proof of Proposition 2.3**. In this proof, we will prove the inequality by seeing it as a $C$-valued to $\ell^2(\mathbb{Z})$-valued inequality and then applying vector-valued Calderón-Zygmund theory to it.

Consider the vector-valued operator $\tilde{S}$ given by

$$\tilde{S} f(x) := (\Delta_j f(x))_{j \in \mathbb{Z}};$$

then, since $\|\tilde{S} f\|_{\ell^2(\mathbb{Z})} = \tilde{S} f$, the inequality we want to prove can be rephrased as

$$\|\tilde{S} f\|_{L^p(\mathbb{R}; \ell^2(\mathbb{Z}))} \lesssim_p \|f\|_{L^p(\mathbb{R})}.$$  \hspace{1cm} (2.4)

The $p = 2$ case is easy to prove: indeed, in this case the left-hand side of (2.4) is simply $\sum_j \|\Delta_j f\|_{L^2(\mathbb{R})}^2$, and by Plancherel’s theorem this is equal to

$$\int_R |\hat{f}(\xi)|^2 \sum_{j \in \mathbb{Z}} |\psi(2^{-j} \xi)|^2 d\xi;$$

the sum is clearly $\lesssim 1$ for any $\xi$ and thus conclude by using Plancherel again.

Next it remains to show $\tilde{S}$ is given by convolution with a vector-valued kernel that satisfies a condition analogous to ii) as in the proof of Proposition 2.1 (we do not need i) and iii) because we have already concluded the $L^2$ boundedness by other means); a final appeal to vector-valued Calderón-Zygmund theory will conclude the proof. It is easy to see that the convolution kernel for $\tilde{S}$ is

$$\tilde{K}(x) := (\hat{\psi}_j(x))_{j \in \mathbb{Z}},$$

where $\hat{\psi}_j(x)$ denotes $2^j \hat{\psi}(2^j x)$. We have to verify the vector-valued Hörmander condition

$$\int_{|x| > 2|y|} \|\tilde{K}(x - y) - \tilde{K}(x)\|_{B(\mathbb{C}, \ell^2(\mathbb{Z}))} dx \lesssim 1.$$
Recall that this is a consequence of the bound $\|\hat{K}'(x)\|_{B(\mathbb{C}, \ell^p(\mathbb{Z}))} \lesssim |x|^{-2}$, where the prime denotes the componentwise derivative. Since $\|K(x)\|_{B(\mathbb{C}, \ell^p(\mathbb{Z}))} = (\sum_j |\hat{\psi}_j(x)|^2)^{1/2}$ (prove this), we have

$$\|\hat{K}'(x)\|^2_{B(\mathbb{C}, \ell^2(\mathbb{Z}))} = \sum_{j \in \mathbb{Z}} 2^{4j} |\hat{\psi}'(2^j x)|^2;$$

thanks to the smoothness of $\psi$, we have $|\hat{\psi}'(x)| \leq (1 + |x|)^{-100}$ (or any other large positive exponent) and the estimate follows easily. We let you fill in the details in Exercise 4.

For the second proof of the proposition, we will introduce a basic but very important tool - Khintchine’s inequality - that allows one to linearise objects of square function type without resorting to duality.

**Lemma 2.4** (Khintchine’s inequality). Let $(\epsilon_k(\omega))_{k \in \mathbb{N}}$ be a sequence of independent identically distributed random variables over a probability space $(\Omega, \mathbb{P})$ taking values in the set \{+1, −1\} and such that each value occurs with probability 1/2. Then for any $0 < p < \infty$ we have that for any sequence of complex numbers $(a_k)_{k \in \mathbb{N}}$

$$\left( \mathbb{E}_\omega \left[ \sum_{k \in \mathbb{N}} |\epsilon_k(\omega)a_k|^p \right] \right)^{1/p} \sim_p \left( \sum_{k \in \mathbb{N}} |a_k|^2 \right)^{1/2},$$

where $\mathbb{E}_\omega$ denotes the expectation with respect to $\mathbb{P}$.

In other words, by randomising the signs the expression $\sum_k \pm a_k$ behaves on average simply like the $\ell^2$ norm of the sequence $(a_k)_k$. See the exercises for some interesting uses of the lemma.

**Proof of Lemma 2.4.** The proof relies on a clever use of the independence assumption.

Let $E_p := \left( \mathbb{E}_\omega \left[ \sum_{k \in \mathbb{N}} |\epsilon_k(\omega)a_k|^p \right] \right)^{1/p}$ for convenience, and observe that $E_2 = \left( \sum_{k \in \mathbb{N}} |a_k|^2 \right)^{1/2}$ (because $\mathbb{E}_\omega[\epsilon_k(\omega)\epsilon_k'(\omega)] = \delta_{k,k'}$; cf. Exercise 1). First we have two trivial facts: when $p > 2$ we have by Hölder’s inequality that $E_2 \leq E_p$, and when $0 < p < 2$ we have conversely that $E_2 \geq E_p$ by Jensen’s inequality. Next, observe that to prove the $\lesssim_p$ part of the inequality it suffices to assume that the $a_k$’s are real-valued. We thus make this assumption and proceed to estimate $E_p$ by expressing the $L^p$ norm as an integral over the superlevel sets,

$$E_p^2 = \int \left| \sum_{k \in \mathbb{N}} \epsilon_k(\omega)a_k \right|^p d\mathbb{P}(\omega) = p \int_0^\infty \lambda^{p-1} \mathbb{P} \left( \sum_{k \in \mathbb{N}} |\epsilon_k(\omega)a_k| > \lambda \right) d\lambda.$$

We split the level set in two and estimate $\mathbb{P} \left( \sum_{k \in \mathbb{N}} \epsilon_k(\omega)a_k > \lambda \right)$; observe that for any $t > 0$ this is equal to $\mathbb{P} \left( e^{t \sum_k \epsilon_k(\omega)a_k} > e^{t \lambda} \right)$, which we estimate by Markov’s inequality by

$$e^{-t\lambda} \int e^{t \sum_k \epsilon_k(\omega)a_k} d\mathbb{P}(\omega).$$

Since $e^{t \sum_k \epsilon_k(\omega)a_k} = \prod_k e^{t \epsilon_k(\omega)a_k}$, it follows by independence of the $\epsilon_k$’s that the integral equals $\prod_k \int e^{t \epsilon_k(\omega)a_k} d\mathbb{P}(\omega)$, and each factor is easily evaluated to be $\cosh(ta_k)$. Since $\cosh(x) \leq e^{x^2/2}$, we have shown that the above is controlled by

$$e^{-t\lambda} \prod_k e^{t^2 a_k^2} = e^{-t\lambda + \frac{1}{2}t^2 \sum_k a_k^2};$$

5Technically, we can only argue so for finite products; but we can assume that at most $N$ coefficients $a_k$ are non-zero and then take the limit $N \to \infty$ at the end of the argument.
if we choose $t = \lambda/\|a\|_2^2$, the above is controlled by $\exp \left(-\frac{\lambda^2}{2\|a\|_2^2}\right)$. Inserting these bounds in the layer-cake representation of $E^p_p$, we have shown that

$$E^p_p \leq 2p \int_0^\infty \lambda^{p-1} e^{-\frac{\lambda^2}{2\|a\|_2^2}} d\lambda = \|a\|_2^{-2p/2} \int_0^\infty s^{p/2-1} e^{-s} ds,$$

which shows $E_p \lesssim_p \|a\|_{L^2(\mathbb{R})} = E_2$ for all $0 < p < \infty$.

Finally, we have to prove $E_2 \lesssim_p E_p$ as well, and by the initial remarks we only need to do so in the regime $0 < p < 2$. For such a $p$, take an exponent $q > 2$ and find $0 < \theta < 1$ such that $\frac{1}{2} = \frac{1-\theta}{p} + \frac{\theta}{q}$; by the logarithmic convexity of the $L^p$ norms we have then $E_2 \leq E_p^{1-\theta} E_q^\theta$, and since we have proven above that $E_q \lesssim_q E_2$ we can conclude.

Now we are ready to provide a second proof of the proposition.

**Second proof of Proposition 2.3.** We claim that it will suffice to prove for every $\omega$ that

$$\left\| \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \tilde{\Delta}_j f \right\|_{L^p(\mathbb{R})} \lesssim_p \|f\|_{L^p}$$

with constant independent of $\omega$. Indeed, if this is the case, then we can raise the above inequality to the exponent $p$ and then apply the expectation $E_\omega$ to both sides. The right-hand side remains $\|f\|_{L^p}$ because it does not depend on $\omega$, and the left-hand side becomes by linearity of expectation

$$\int_{\mathbb{R}} E_\omega \left[ \left| \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \tilde{\Delta}_j f(x) \right|^p \right] dx,$$

which by Khintchine’s inequality is comparable to $\int |\tilde{\mathcal{S}}f(x)|^p dx$, thus proving the claim.

To prove the $L^p$ boundedness of the operators $\tilde{T}_\omega f(x) := \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \tilde{\Delta}_j f$ we appeal to a result encountered in a previous lecture, namely the Hörmander-Mikhlin multiplier theorem. Indeed, $\tilde{T}_\omega$ is given by convolution with the kernel $K_\omega := \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \hat{\psi}_j$, whose Fourier transform is simply

$$K_\omega(\xi) = \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \hat{\psi}(2^{-j} \xi).$$

For any $\xi$, there are at most 3 terms in the sum above that are non-zero; this and the other assumptions on $\psi$ readily imply that, uniformly in $\omega$,

$$|K_\omega(\xi)| \lesssim 1,$$

$$\left| \frac{d}{d\xi} K_\omega(\xi) \right| \lesssim |\xi|^{-1};$$

therefore $K_\omega$ is indeed a Hörmander-Mikhlin multiplier, and by the Hörmander-Mikhlin multiplier theorem we have that $\tilde{T}_\omega$ is $L^p \to L^p$ bounded for $1 < p < \infty$ with constant uniform in $\omega$, and we are done.

**2.3. Littlewood-Paley square function.** With the material developed in the previous subsections we are now ready to prove (†). We restate it properly:

**Theorem 2.5.** Let $1 < p < \infty$ and let $Sf$ denote the Littlewood-Paley square function

$$Sf := \left( \sum_j |\Delta_j f|^2 \right)^{1/2}.$$

Then for all functions $f \in L^p(\mathbb{R})$ it holds that

$$\|Sf\|_{L^p(\mathbb{R})} \sim_p \|f\|_{L^p(\mathbb{R})}.$$

(†)
Proposition 2.6. Let \( \|f\|_{L^p} = \sup_{g \in L^{p'}; \|g\|_{p'} = 1} \int fg \, dx \),

we can write

\[
\int fg \, dx = \int \left( \sum_j \Delta_j f \right) \left( \sum_k \Delta_k g \right) \, dx = \sum_j \int \Delta_j f \Delta_k g \, dx = \int \sum_j \Delta_j f \Delta_j g \, dx
\]

\[
\leq \int \left( \sum_j |\Delta_j f|^2 \right)^{1/2} \left( \sum_j |\Delta_j g|^2 \right)^{1/2} \, dx = \int Sf Sg \, dx,
\]

where we have used the orthogonality of the projections in the third equality and then Cauchy-Schwarz inequality. Now Hölder’s inequality shows that

\[
\int fg \, dx \leq \|Sf\|_{L^p} \|Sg\|_{L^{p'}},
\]

and by the assumed \( L^{p'} \to L^{p'} \) boundedness of \( S \) we can bound the right hand side by \( \lesssim_p \|Sf\|_{L^p}\|g\|_{L^{p'}} = \|Sf\|_{L^p} \), thus concluding that \( \|f\|_{L^p} \lesssim_p \|Sf\|_{L^p} \) as well. (See Exercise 11 for why going the opposite direction would not work).

Now observe the following fundamental fact: if \( \Delta_j \) denotes the smooth frequency projection as defined in Section 2.2, then we have for any \( j \)

\[
\Delta_j \Delta_j = \Delta_j;
\]

indeed, recall that \( \Delta_j f(\xi) = \psi(2^{-j} \xi) \tilde{f}(\xi) \) and that \( \psi(2^{-j} \xi) \) is identically equal to 1 on the intervals \( [2^j, 2^{j+1}] \cup [-2^{j+1}, -2^j] \).

Using this fact, we can argue by Corollary 2.2 that

\[
\left\| \left( \sum_j |\Delta_j f|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R})} = \left\| \left( \sum_j |\Delta_j \Delta_j f|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R})} \lesssim_p \left\| \left( \sum_j |\Delta_j f|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R})};
\]

but the right-hand side is now the smooth Littlewood-Paley square function \( \tilde{S}f \), and by Proposition 2.3 it is bounded by \( \lesssim_p \|f\|_{L^p} \), thus concluding the proof. \( \square \)

2.4. Higher dimensional variants. So far we have essentially worked only in dimension \( d = 1 \) (except for Proposition 2.1). There are however some generalisations to higher dimensions of the theorems above, whose proofs follow from similar arguments.

Let \( d > 1 \). First of all, with \( \psi \) as in Section 2.2 define the (smooth) annular frequency projection\( ^6 \) \( \tilde{P}_j \) by

\[
\tilde{P}_j f(\xi) := \psi(2^{-j} \xi) \tilde{f}(\xi).
\]

Then we have that the corresponding square function satisfies the analogue of Proposition 2.3.

**Proposition 2.6.** Let \( \tilde{S} \) denote the square function

\[
\tilde{S}f = \left( \sum_{j \in \mathbb{Z}} |\tilde{P}_j f|^2 \right)^{1/2}.
\]

Then for any \( 1 < p < \infty \) we have for any \( f \in L^p(\mathbb{R}^d) \) the inequality

\[
\|\tilde{S}f\|_{L^p(\mathbb{R}^d)} \lesssim_{p,d} \|f\|_{L^p(\mathbb{R}^d)}.
\]

\(^6\) We call them “annular” because \( \psi(2^{-j} \xi) \) is smoothly supported in the annulus \( \{ \xi \in \mathbb{R}^d : 2^{j-1} \leq |\xi| \leq 2^{j+1} \} \).
Either of the proofs given for Proposition 2.3 generalises effortlessly to this case. Interestingly, the non-smooth analogue of $\tilde{S}$ does not satisfy the analogue of Theorem 2.5. Indeed, the reason behind this fact is that the operator taking on the rôle of the Hilbert transform $H$ would be the so-called ball multiplier, given by

$$T_B f(x) := \int_{\mathbb{R}^d} 1_{B(0,1)}(\xi) \hat{f}(\xi)e^{2\pi i \xi \cdot x} d\xi;$$  \hspace{1cm} (2.6)

however, as seen before, it is a celebrated result of Charles Fefferman that the operator $T_B$ is only bounded when $p = 2$, and unbounded otherwise. This result is very deep and we do not have enough room to discuss it in these notes. One of the ingredients needed in the proof is nevertheless presented in Exercise 10, if you are interested.

Another way in which one can generalise Littlewood-Paley theory to higher dimensions is to take products of the dyadic intervals $[2^k, 2^{k+1}]$. That is, let for convenience $I_k$ denote the dyadic Littlewood-Paley interval

$$I_k := [2^k, 2^{k+1}] \cup [-2^k, -2^{k+1}];$$

then for $\underline{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d$ one defines the “rectangle” $R_{\underline{k}}$ to be

$$R_{\underline{k}} = I_{k_1} \times \ldots \times I_{k_d}$$

and defines the square function $S_{\text{rect}}$ to be

$$S_{\text{rect}} f := \left( \sum_{\underline{k} \in \mathbb{Z}^d} |\Delta_{R_{\underline{k}}} f|^2 \right)^{1/2}.$$

It is easy to see that the above rectangles are all disjoint and they tile $\mathbb{R}^d$. Then we have the rectangular analogue of Theorem 2.5

**Theorem 2.7.** For any $1 < p < \infty$ and all functions $f \in L^p(\mathbb{R}^d)$ it holds that

$$\|S_{\text{rect}} f\|_{L^p(\mathbb{R}^d)} \sim_{p,d} \|f\|_{L^p(\mathbb{R}^d)}.$$

**Proof.** We will deduce the theorem from the one-dimensional case, that is from Theorem 2.5. With the same argument given there, one can see that it will suffice to establish the $\lesssim_{p,d}$ part of the inequalities.

First of all, with the same notation as in Lemma 2.4 when $d = 1$ we define the operators

$$T_\omega f := \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \Delta_j f.$$  

These operators are $L^p(\mathbb{R}) \to L^p(\mathbb{R})$ bounded for any $1 < p < \infty$ with constant independent of $\omega$, as a consequence of Theorem 2.5 (prove this in Exercise 12). Next, when $d > 1$, define $\epsilon_{\underline{k}}(\omega) := \epsilon_{k_1}(\omega) \cdot \ldots \cdot \epsilon_{k_d}(\omega)$. By Khintchine’s inequality, it will suffice to prove that the operator

$$T_{\omega} f := \sum_{\underline{k} \in \mathbb{Z}^d} \epsilon_{\underline{k}}(\omega) \Delta_{R_{\underline{k}}} f$$

is $L^p \to L^p$ bounded independently of $\omega$ (you are invited to check that this is indeed enough). However, it is easy to see that

$$T_{\omega} = T_{\omega}^{(1)} \circ \ldots \circ T_{\omega}^{(d)},$$

where $T_{\omega}^{(k)}$ denotes the operator $T_{\omega}$ above applied to the $x_k$ variable. The boundedness of $T_{\omega}$ thus follows from that of $T_{\omega}$ by integrating in one variable at a time. \qed
3. Applications

In this section we will present two applications of the Littlewood-Paley theory developed so far. You can find further applications in the exercises (see particularly Exercise 22 and Exercise 23).

3.1. Marcinkiewicz multipliers. Given an $L^\infty(\mathbb{R}^d)$ function $m$, one can define the operator $T_m$ given by

$$T_m f(\xi) := m(\xi) \hat{f}(\xi)$$

for all $f \in L^2(\mathbb{R}^d)$. The operator $T_m$ is called a multiplier and the function $m$ is called the symbol of the multiplier. Since $m \in L^\infty$, Plancherel’s theorem shows that $T_m$ is a linear operator bounded in $L^2$; its definition can then be extended to $L^2 \cap L^p$ functions (which are dense in $L^p$). A natural question to ask is: for which values of $p$ in $1 \leq p \leq \infty$ is the operator $T_m$ an $L^p \to L^p$ bounded operator? When $T_m$ is bounded in a certain $L^p$ space, we say that it is an $L^p$-multiplier.

The operator $T_m$ introduced in Section 1 is an example of a multiplier, with symbol $m(\xi, \tau) = \tau/(\tau - 2\pi i |\xi|^2)$. We have seen that it cannot be a (euclidean) Calderón-Zygmund operator, and thus in particular it cannot be a Hörmander-Mikhlin multiplier. This can be seen more directly by the fact that any Hörmander-Mikhlin condition of the form $|\partial^\alpha m(\xi, \tau)| \lesssim |(\xi, \tau)|^{-|\alpha|} = (|\xi|^4 + \tau^2)^{-|\alpha|/2}$ is clearly incompatible with the rescaling invariance of the symbol $m$, which satisfies $m(\lambda^2 \xi, \lambda^2 \tau) = m(\xi, \tau)$ for any $\lambda \neq 0$. However, the derivatives of $m$ actually satisfy some other superficially similar conditions that are of interest to us. Indeed, letting $(\xi, \tau) \in \mathbb{R}^d$ for simplicity, we can see for example that $\partial_\xi \partial_\tau m(\xi, \tau) = \lambda^4 \partial_\xi \partial_\tau m(\lambda^2 \xi, \lambda^2 \tau)$. When $|\tau| \leq |\xi|^2$ we can therefore argue that $|\partial_\xi \partial_\tau m(\xi, \tau)| = |\tau|^{-3} |\partial_\xi \partial_\tau m(1, \tau)| |\xi|^{-2} \lesssim |\xi|^{-1} |\tau|^{-1} \sup_{|\eta| \leq 1} |\partial_\xi \partial_\tau m(1, \eta)|$, and similarly when $|\tau| \geq |\xi|^2$; this shows that for any $(\xi, \tau)$ with $\xi, \tau \neq 0$ one has

$$|\partial_\xi \partial_\tau m(\xi, \tau)| \lesssim |\tau|^{-1}. \tag{3.1}$$

This condition is comparable with the corresponding Hörmander-Mikhlin condition only when $|\xi| \sim |\tau|$, and is vastly different otherwise, being of product type (also notice that the inequality above is compatible with the rescaling invariance of $m$, as it should be).

The Littlewood-Paley theory developed in these notes allows us to treat multipliers with symbols that satisfy product-type conditions like the above, and which typically are beyond the reach of Calderón-Zygmund theory. Indeed, we have the following general result, where the conditions assumed of the multiplier are a generalisation of the pointwise ones above.

**Theorem 3.1 (Marcinkiewicz multiplier theorem).** Let $m$ be a function of $\mathbb{R}^d$ that is of class $C^d$ away from the hyperplanes where one of the coordinates is zero. Suppose that $m$ satisfies the following conditions:

i) $m \in L^\infty$;

ii) there is a constant $C > 0$ such that for every $0 < \ell \leq d$ and for every permutation $j_1, \ldots, j_\ell, j_{\ell+1}, \ldots, j_d$ of the set $\{1, \ldots, d\}$ we have

$$\sup_{\xi, \eta \in \mathbb{R}^d} \sup_{\xi_{\ell+1} \neq 0} \int_{R_{\xi, \eta}} |\partial_{j_1} \cdots \partial_{j_{\ell}} m(\xi)| d\xi_{j_1} \cdots d\xi_{j_\ell} \leq C,$$

where the $R_{\xi, \eta}$ are the dyadic Littlewood-Paley rectangles as in §2.4.

Then the multiplier $T_m$ associated to the symbol $m$ satisfies for any $1 < p < \infty$ the inequality

$$\|T_m f\|_{L^p(\mathbb{R}^d)} \lesssim_p \|f\|_{L^p(\mathbb{R}^d)}.$$
A multiplier whose symbol satisfies the conditions of the above theorem is called a Marcinkiewicz multiplier. Condition ii), spelled out in words, says that a certain subset of the partial derivatives \( \partial^\alpha m \) of \( m \) (precisely, the derivatives \( \partial^\alpha \) where the components of the multi-index \( \alpha \) have values in \( \{0,1\} \) and \( 0 < |\alpha| \leq d \) has the property that their integral over any \( |\alpha| \)-dimensional dyadic Littlewood-Paley rectangle is uniformly bounded, where the integration is happening in those variables \( \xi_j \) such that \( \alpha_j \neq 0 \).

**Remark 2.** The conditions above are very general but also very cumbersome to spell out, as the above attempt demonstrates. The pointwise conditions, analogous to the one seen before, that imply condition ii) of the theorem are usually easier to check and are as follows:

\[ ii') \text{ for any multi-index } \alpha \in \{0,1\}^d \text{ such that } 0 < |\alpha| \leq d \text{ we have} \]
\[ |\partial^\alpha m(\xi)| \lesssim |\xi_1|^{-\alpha_1} \cdots |\xi_d|^{-\alpha_d}. \]

See Exercise 14.

When the dimension \( d \) is equal to 1, the statement of the theorem can be superficially generalised and reformulated in the following way, that perhaps clarifies the moral meaning of condition ii).

**Theorem 3.2** (Marcinkiewicz multiplier theorem for \( d = 1 \)). Let \( m \) be a function of \( \mathbb{R} \) that is of bounded variation on any interval not containing the origin. Suppose that \( m \) satisfies the following conditions:

1) \( m \in L^\infty; \)

2) there is a constant \( C > 0 \) such that, with \( \int_I |dm| \) the total variation of \( m \) over the interval \( I \),

\[ \sup_{k \in \mathbb{Z}} \int_{[2^k, 2^{k+1}) \cup [-2^k, -2^{k+1})} |dm|(\xi) \leq C. \]

Then the multiplier \( T_m \) with symbol \( m \) satisfies for any \( 1 < p < \infty \) the inequality

\[ \|T_m f\|_{L^p(\mathbb{R})} \lesssim_p \|f\|_{L^p(\mathbb{R})}. \]

Condition ii) is now saying that the total variation of \( m \) on any dyadic Littlewood-Paley interval is bounded. Observe that it is implied by the pointwise condition \( |m'(\xi)| \lesssim |\xi|^{-1}, \) but since we are in dimension \( d = 1 \) now this stronger condition would just coincide with the Hörmander-Mikhlin one. There is in general a certain degree of overlap between the Marcinkiewicz multiplier theorem and the Hörmander-Mikhlin theorem in any dimension, but neither implies the other.

We will prove Theorem 3.2 here, and the proof we will give will already contain all the main ingredients needed for the proof of the full Theorem 3.1. We leave it to you to extend the proof to higher dimensions in Exercise 15.

**Proof of Theorem 3.2.** Let \( T = T_m \) be the multiplier with symbol \( m \). Observe that by (Theorem 2.5) we have

\[ \|T f\|_{L^p} \lesssim_p \left\| \left( \sum_{j \in \mathbb{Z}} |\Delta_j T f|^2 \right)^{1/2} \right\|_{L^p}, \]

and thus it will suffice to prove that

\[ \left\| \left( \sum_{j \in \mathbb{Z}} |\Delta_j T f|^2 \right)^{1/2} \right\|_{L^p} \lesssim_p \left\| \left( \sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \right)^{1/2} \right\|_{L^p} \tag{3.1} \]

\[ \text{Recall that the total variation of a function } f \text{ on the interval } [a, b] \text{ is defined as } \sup_N \sup_{\xi_0 \leq \xi_1 \leq \cdots \leq \xi_N \leq b} \sum_{j=1}^{N-1} |f(\xi_{j+1}) - f(\xi_j)|; \text{ if } f \text{ has finite total variation on } [a, b] \text{ then there exists a complex Borel measure } \mu \text{ such that } f(x) = f(a) + \int_a^x d\mu. \]
(again by \(\text{[1]}\)). Now, \(\Delta_j T\) is a multiplier with symbol \(m(\xi)1_{[2^j, 2^{j+1}]}(\xi)\) (we discard the negative frequencies for ease of notation), and by condition 2) we see that there exists a complex Borel measure \(\text{[1]}\) such that for \(\xi \in [2^j, 2^{j+1}]\)

\[
m(\xi) = m(2^j) + \int_2^\xi dm(\eta).
\]

By Fubini we see therefore that

\[
\Delta_j T f(x) = m(2^j) \Delta_j f(x) + \int_{2^j}^{2^{j+1}} \Delta_{[\eta, 2^{j+1}]} f(x) dm(\eta),
\]

where we recall that \(\Delta_{[\eta, 2^{j+1}]}\) denotes frequency projection onto the interval \([\eta, 2^{j+1}]\). By condition 1), the first term in the right-hand side contributes at most \(\|m\|_{L^\infty} S f\) to the left-hand side of (3.1) overall, and therefore we can safely discard it; let \(T_j f(x)\) be the second term. We have by Cauchy-Schwarz and condition 2) that

\[
|T_j f(x)|^2 \leq C \int_{2^j}^{2^{j+1}} |\Delta_{[\eta, 2^{j+1}]} f(x)|^2 dm(\eta),
\]

and therefore we have reduced to control the square function

\[
\left( \sum_{j \in \mathbb{Z}} \int_{2^j}^{2^{j+1}} |\Delta_{[\eta, 2^{j+1}]} f(x)|^2 dm(\eta) \right)^{1/2}.
\]

Now comes the tricky part: we want to realise the above expression as the norm in an abstract Hilbert space \(\mathcal{H}\) of a vector-valued frequency projection, as in Proposition 2.1. There are many ways of doing so, and the following is the one we choose. Let \(\Gamma := \bigcup_{j \in \mathbb{Z}} (2^j, 2^{j+1}] \times \{j\}\), that is the set of elements \((\eta, j)\) such that \(2^j < \eta \leq 2^{j+1}\); observe that a set \(E \subseteq \Gamma\) has necessarily the form \(E = \bigcup_{j \in A} E_j \times \{j\}\) for a certain \(A \subseteq \mathbb{Z}\) and sets \(E_j \subseteq (2^j, 2^{j+1}]\). We define the measure \(d\mu\) on \(\Gamma\) to be given by

\[
\mu(E) := \sum_{j \in A} \int_{E_j} dm(\eta),
\]

provided the \(E_j\)'s are \(|dm|\)-measurable. Thus the Hilbert space \(\mathcal{H} = L^2(\Gamma, d\mu)\) consists of those functions \(G(\eta, j)\) such that

\[
\sum_{j \in \mathbb{Z}} \int_{2^j}^{2^{j+1}} |G(\eta, j)|^2 dm(\eta)
\]

is finite. Finally, to any \((\eta, j) \in \Gamma\) we assign the interval \(I_{\eta, j} := [\eta, 2^{j+1}]\). With these definitions, we see that (3.2) corresponds to

\[
\|(\Delta_{I_{\eta, j}} g_{\eta, j}(x))(\eta, j)\|_{\mathcal{H}},
\]

where in our particular case we can take \(g_{\eta, j}(x) = \Delta_j f(x)\) because \(\Delta_{I_{\eta, j}} \Delta_j = \Delta_{I_{\eta, j}}\). By Proposition 2.1 we have for generic vector-valued functions \((g_{\eta, j})(\eta, j) \in \Gamma\) that

\[
\|(\Delta_{I_{\eta, j}} g_{\eta, j})(\eta, j)\|_{L^p(\mathbb{R}, \mathcal{H})} \lesssim_p \|(g_{\eta, j})(\eta, j)\|_{L^p(\mathbb{R}, \mathcal{H})};
\]

applying this to \(g_{\eta, j} = \Delta_j f\) we see that we have bounded \(\|(\sum_j |T_j f|^2)^{1/2}\|_{L^p}\) by the \(L^p(\mathbb{R})\) norm of

\[
\left( \sum_{j \in \mathbb{Z}} \int_{2^j}^{2^{j+1}} |\Delta_j f(x)|^2 dm(\eta) \right)^{1/2},
\]

but the integrand does not depend on \(\eta\) and thus condition 2) shows that the above is bounded pointwise by \(C S f(x)\). This concludes the proof of (3.1) and thus the proof of the theorem. \(\square\)
3.2. Boundedness of the spherical maximal function. Recall that the boundedness of the spherical maximal function \( \mathcal{M}_{S^d} \) implies, through the method of rotations, that the Hardy-Littlewood maximal function \( M \) is \( L^q(\mathbb{R}^d) \rightarrow L^q(\mathbb{R}^d) \) bounded for any \( 1 < q < \infty \) with constant independent of the dimension. In the final part of this lecture we will finally prove the boundedness of \( \mathcal{M}_{S^d} \) (when \( d \geq 3 \)).

Theorem 3.3. Let \( d \geq 3 \). Then for any \( \frac{d}{d-1} < p \leq \infty \) the inequality
\[
\| \mathcal{M}_{S^d} f \|_{L^p(\mathbb{R}^d)} \lesssim_p \| f \|_{L^p(\mathbb{R}^d)}
\]
holds for any \( f \in L^p(\mathbb{R}^d) \).

The range of \( p \)'s stated in the above theorem is sharp (prove it in Exercise 21). A small caveat: the proof we will give below will yield a constant \( C_{d,p} \) for the above inequality that depends on the dimension \( d \). It is only later - through the method of rotations - that one can remove this dependence on the dimension, showing that if one has a bound \( \| \mathcal{M}_{S^d} f \|_{L^p(\mathbb{R}^d)} \rightarrow L^p(\mathbb{R}^d) \leq A \) for dimension \( d \) then one also has the bound \( \| \mathcal{M}_{S^d} f \|_{L^p(\mathbb{R}^{d+1})} \rightarrow L^p(\mathbb{R}^{d+1}) \leq A \) for dimension \( d + 1 \).

Proof. It will suffice to prove the theorem for \( \frac{d}{d-1} < p < 2 \), since the rest of the exponents can be obtained by Marcinkiewicz interpolation with the trivial \( L^\infty \) estimate. It will also suffice to assume that \( f \) is a Schwartz function for convenience, as they are dense in \( L^p \).

We will prove the theorem first for a local version of \( \mathcal{M}_{S^d} \), namely the operator
\[
\mathcal{M}^{\text{loc}}_{S^d} f(x) := \sup_{1 \leq r \leq 2} |A_r f(x)|,
\]
where \( A_r \) denotes the average over the sphere of radius \( r \), that is
\[
A_r f(x) := \int_{S^d} f(x - ry) d\sigma(y),
\]
where \( d\sigma \) is the normalised surface measure on the sphere. Then in Exercise 21 you will conclude the proof for the full case of \( \mathcal{M}_{S^d} \) with little extra effort.

We will use an annular frequency decomposition as the one in Section 2.4, but slightly modified to suit our purposes. We let \( \psi \) be a smooth radial function compactly supported in the annulus \( \{ \xi \in \mathbb{R}^d : 1/2 < |\xi| < 2 \} \) and such that \( \sum_{\xi \in \mathbb{Z}} \psi(2^{-j} \xi) = 1 \). We let \( \psi_j(\xi) := \psi(2^{-j} \xi) \) for convenience. Thus, with \( P_j \) being given by \( P_j f(\xi) = \psi_j(\xi) \hat{f}(\xi) \) we have that for all functions \( f \in L^2 \) we can decompose
\[
f = \sum_{j \in \mathbb{Z}} P_j f.
\]
Since the radius \( r \) will be \( \sim 1 \), we will not need to consider the projections to extremely low frequencies separately; therefore we define \( P_{\text{low}} f := \sum_{j<0} P_j f \), so that \( f = P_{\text{low}} f + \sum_{j \in \mathbb{N}} P_j f \). In view of the above decomposition, to conclude the \( L^p \rightarrow L^p \) boundedness of \( \mathcal{M}_{S^d} \), it will suffice by triangle inequality to prove
\[
\| \sup_{1 \leq r \leq 2} |A_r P_{\text{low}} f| \|_{L^p} \lesssim_p \| f \|_{L^p},\tag{3.3}
\]
\[
\| \sup_{1 \leq r \leq 2} |A_r P_j f| \|_{L^p} \lesssim_p 2^{-\delta j} \| f \|_{L^p},\tag{3.4}
\]
for some \( \delta = \delta_{p,d} > 0 \).

Now, the heuristic motivation behind such a decomposition is that \( |P_j f| \) is now\footnote{Recall that such a function exists, and can be realised for example by taking \( \psi(\xi) = \varphi(\xi) - \varphi(\xi/2) \) with \( \varphi \) smooth compactly supported in \([-2,2] \) and identically 1 on \([-1,1] \).}
roughly constant at scale $2^{-j}$. Indeed, one way to appreciate this informal principle (which is a manifestation of the Uncertainty Principle) is to show for example that $|P_j f(x)|$ is pointwise dominated by its average at scale $2^{-j}$: since $\hat{P_j f}$ is supported in the ball $B(0, 2^{j+1})$, we can take a smooth bump function $\phi$ such that $\hat{\phi}(2^{-j} \xi)$ is identically 1 on this ball and vanishes outside $B(0, 2^{j+2})$ and see that therefore $\hat{P_j f}(\xi) = \phi(2^{-j} \xi) \hat{P_j f}(\xi)$. The latter means that $P_j f = P_j f \ast 2^j \hat{\phi}(2^j \cdot)$ and the smoothness of $\phi$ means that we have $|\hat{\phi}(x)| \lesssim (1 + |x|)^{-10d}$ (say); therefore we have by expanding the convolution that

$$|P_j f(x)| \lesssim \int_{\mathbb{R}^d} |P_j f(x - 2^{-j} y)| \frac{dy}{(1 + |y|)^{10d}}.$$ 

This observation suggests that $A_r P_{\text{low}} f$ should just be an average of $f$ at unit scale, since $r \sim 1$ and $P_{\text{low}} f$ is roughly constant at scales $\lesssim 1$. Indeed, one can realise that $P_{\text{low}}$ is given by convolution with a Schwartz function $\varphi$ and thus by Fubini $A_r P_{\text{low}} f = f \ast A_r \varphi$; it is not hard to show that $\varphi$ being Schwartz and $r$ being $\sim 1$ implies bounds such as $|A_r \varphi(x)| \lesssim (1 + |x|)^{-10d}$ (or any other exponent really), and therefore we have as seen before

$$|A_r P_{\text{low}} f(x)| \lesssim M f(x)$$

uniformly in $r$, so that the pointwise domination continues to hold after we take the supremum. This immediately implies (3.3) for any $1 < p < \infty$ (so even outside our desired range).

Repeating the argument for $P_j f$ does not allow us to conclude. Indeed, you can see that in this case $|A_r \hat{\psi_j}|$ is essentially a function of magnitude $\sim 2^j$ concentrated in a shell of width $\sim 2^{-j}$ and radius $r$ (see Exercise 20). A greedy argument, disregarding the width information, shows the rough bound $|A_r \hat{\psi}_j(x)| \lesssim 2^j (1 + |x|)^{-10d}$, which in turn implies $|A_r P_j f| \lesssim 2^j M f$ uniformly in $r \sim 1$. Therefore this argument gives at best the bound

$$\left\| \sup_{1 \leq r \leq 2} |A_r P_j f| \right\|_{L^p} \lesssim 2^j \|f\|_{L^p}, \quad (3.5)$$

which blows up as $j \to \infty$. However, one advantage of this bound is that it holds even for $1 < p \leq d/(d-1)$, which would allow for some interpolation if we had a decaying bound for some other exponent to compensate with. A particularly nice exponent is $p = 2$ of course, because Plancherel’s theorem is available, so we should explore what happens in this case. Consider $r$ fixed and observe that

$$\hat{A_r f}(\xi) = \hat{d\sigma}(r \xi) \hat{f}(\xi),$$

where $\hat{d\sigma}$ denotes the Fourier transform of $d\sigma$, that is $\hat{d\sigma}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i y \cdot \xi} d\sigma(y)$ (verify the formula). It turns out that, due to the curvature of the sphere, $\hat{d\sigma}$ has a certain amount of decay as $\xi$ gets large - in particular, as you will see in a future lecture, one has the pointwise bound

$$|\hat{d\sigma}(\xi)| \lesssim (1 + |\xi|)^{-(d-1)/2}.$$ 

We will take this for granted here, but if you cannot wait you can prove the decay in Exercise 19. Now, combining the decay information above with the frequency localisation of $P_j f$, we see by Plancherel that, since $r \sim 1$,

$$\|A_r P_j f\|_{L^2} \lesssim 2^{-j(d-1)/2} \|f\|_{L^2}, \quad (3.6)$$

a bound nicely decaying in $j$. However, this bound is not directly useful to us as we still have to take the supremum in $r$! To deal with this issue we will do something
a bit unconventional: since for a generic function $\phi(r)$ we have by the Fundamental
Theorem of Calculus that
\[
\sup_{1 \leq r \leq 2} |\phi(r)| \leq |\phi(1)| + \int_1^2 |\phi'(r)|dr,
\]
we would like to replace the supremum by the right-hand side. This is usually a
good idea when the supremum is taken over a controlled interval and the derivative
is a nice object (as will be the case for us); however, if we proceeded with this
approach we would get bad bounds in the end - the second term at the right-hand
side would contribute much more than the other one. We need therefore some more
sophisticated inequality which allows us to optimally balance the two contributions,
and the following will do. Observe that, by the Fundamental Theorem of Calculus
and Cauchy-Schwarz, if $\phi(1) = 0$ we have
\[
\frac{1}{2} \sup_{1 \leq r \leq 2} |\phi(r)|^2 \leq \int_1^2 |\phi(r)||\phi'(r)|dr
\leq \left(\int_1^2 |\phi(r)|^2dr\right)^{1/2} \left(\int_1^2 |\phi'(r)|^2dr\right)^{1/2},
\]
and by the trivial inequality $xy \leq Bx^2 + B^{-1}y^2$ we have therefore
\[
\sup_{1 \leq r \leq 2} |\phi(r)| \leq B\left(\int_1^2 |\phi(r)|^2dr\right)^{1/2} + B^{-1} \left(\int_1^2 |\phi'(r)|^2dr\right)^{1/2}
\]
for any $B > 0$. We take $\phi(r) = A_r P_j f - A_r P_j f$ and estimate the terms on the
right-hand side separately. The $L^2$ bound (3.6) shows that, since the $L^2$ expressions
commute,
\[
\left\|\left(\int_1^2 |A_r P_j f(x)|^2 dx\right)^{1/2}\right\|_{L^2} \lesssim 2^{-j(d-1)/2}\|f\|_{L^2}
\]
(and similarly for $A_1 P_j f$). For the other term, we need to evaluate $\frac{d}{dr} A_r P_j f$ (the
$A_1 P_j f$ term disappears), and this is easy on the Fourier side, giving
\[
\frac{d}{dr} A_r P_j f(\xi) = \xi \cdot \nabla \hat{d}\sigma(r\xi) \hat{P}_j f(\xi).
\]
The gradient $\nabla \hat{d}\sigma$ has the same decay as $\hat{d}\sigma$, namely
$|\nabla \hat{d}\sigma(\xi)| \lesssim (1 + |\xi|)^{-(d-1)/2}$
(and this can be proven in the same way). Combining this decay information with
the frequency localisation of $P_j f$ we obtain by Plancherel’s theorem that
\[
\left\|\left(\int_1^2 \left|\frac{d}{dr} A_r P_j f\right|^2 dr\right)^{1/2}\right\|_{L^2} \lesssim 2^{2j(d-1)/2}\|f\|_{L^2} = 2^{-j(d-1)/2}\|f\|_{L^2}.
\]
Putting all this information together, we see that inequality (3.7) gives us a bound of
$(B2^{-j(d-1)/2} + B^{-1}2^{-j(d-3)/2})\|f\|_{L^2}$ for $\|\sup_{1 \leq r \leq 2}|A_r P_j f\|_{L^2}$, which is optimised
if we choose $B = 2^{1/2}$, resulting in
\[
\|\sup_{1 \leq r \leq 2} |A_r P_j f|\|_{L^2} \lesssim 2^{-j(d-2)/2}\|f\|_{L^2}
\]
(observe that when $d = 2$ this expression gives no decay at all, thus explaining
why the proof does not work in that case). For a given $1 < p < 2$ we can then
interpolate between estimates (3.5) and (3.8) to obtain a constant for the $L^p \to L^p$
norm of $\sup_{1 \leq r \leq 2} |A_r P_j f|$ that is at most $\lesssim 2^{-j(d-1-d/p)+\epsilon}$ for any small $\epsilon > 0$
(do the calculation; you will need to interpolate between the $p = 2$ exponent and an
exponent extremely close to 1 but not 1 - hence the $\epsilon$). The expression $d - 1 - d/p$
is only positive when $p > d/(d-1)$, and therefore in this regime we have that (3.4)
holds for some $\delta = \delta_{p,d} > 0$, allowing us to sum in $j \in \mathbb{N}$ and thus to conclude that
$\mathcal{M}_{g^{loc}}^{\infty}$ is bounded. \qed
Exercises

Exercises that require a bit more work are labeled by a ★.

Exercise 1. Let \((\epsilon_k)_{k \in \mathbb{N}}\) be independent random variables taking values in \(\{+1, -1\}\) such that \(\epsilon_k\) takes each value with equal probability. Show that
\[
E \left[ \left| \sum_{k=1}^{N} \epsilon_k \right|^2 \right] = N.
\]
Generalise this to the case where the sum is \(\sum_{k=1}^{N} \epsilon_k a_k\) with \(a_k\) some complex coefficients.

Exercise 2. Show that (2.2) holds for any collection \(\mathcal{I}\) of disjoint intervals.

Exercise 3. Show that Corollary 2.2 is a special case of Proposition 2.1. What is the measure space \((\Gamma, d\mu)\) that gives the corollary? Go to the proof of the proposition and spell out all the Hilbert norms in the argument in terms of the measure space you found.

Exercise 4. Fill in the missing details in the first proof of Proposition 2.3 (that is, the pointwise bound for \(\|K'(x)\|_{B(\mathbb{C}, \ell^2(\mathbb{Z}))}\)).

Exercise 5. Look at the proof of Khintchine’s inequality (Lemma 2.4). What is the asymptotic behaviour as \(p \to \infty\) of the constant \(C_p\) in the inequality
\[
\left( E_{\omega} \left[ \left| \sum_{k \in \mathbb{N}} \epsilon_k(\omega) a_k \right|^p \right] \right)^{1/p} \lesssim C_p \left( \sum_{k \in \mathbb{N}} |a_k|^2 \right)^{1/2}
\]
obtained in that proof?

Exercise 6. Let \(T\) be a linear operator that is \(L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)\) bounded for some \(1 \leq p < \infty\). Show, using Khintchine’s inequality and the linearity of expectation, that for any vector-valued function \((f_j)_{j \in \mathbb{Z}}\) in \(L^p(\mathbb{R}^d; \ell^2(\mathbb{Z}))\) we have
\[
\left\| \left( \sum_{j \in \mathbb{Z}} |Tf_j|^2 \right)^{1/2} \right\|_{L_p(\mathbb{R}^d)} \lesssim \left\| \left( \sum_{j \in \mathbb{Z}} |f_j|^2 \right)^{1/2} \right\|_{L_p(\mathbb{R}^d)}.
\]
In other words, Khintchine’s inequality provides us with an upgrade to vector-valued estimates, free of charge.

Exercise 7. Recall that the Hausdorff-Young inequality says that for any \(1 \leq p \leq 2\) it holds that
\[
\|\hat{f}\|_{L^{p'}} \leq \|f\|_{L^p}.
\]
Crucially, the inequality cannot hold for \(p > 2\) (notice that when \(p > 2\) one has \(p > 2 > p'\)) and this can be seen in many ways. In this exercise you will show this fact using Khintchine’s inequality - this is an example of the technique known as randomisation, which is useful to construct counterexamples.

i) Let \(\varphi\) be a smooth function with compact support in the unit ball \(B(0, 1)\) and let \(N > 0\) be an integer. Choose vectors \(x_j \in \mathbb{R}^d\) for \(j = 1, \ldots, N\) such that the translated functions \((\varphi(\cdot - x_j))_{j=1,\ldots,N}\) have pairwise disjoint supports and define the function
\[
\Phi_\omega(x) := \sum_{j=1}^{N} \epsilon_j(\omega) \varphi(x - x_j).
\]
Show that
\[
\|\Phi_\omega\|_{L^p} \sim N^{1/p}
\]
for any \(\omega\).
ii) Show, using Khintchine’s inequality, that
\[ \mathbb{E}_\omega \left[ \left\| \hat{\Phi}_\omega \right\|_{L^p}^p \right] \lesssim N^{p/2}. \]

iii) Deduce that there is an \( \omega \) (equivalently, a choice of signs in \( \sum_{j=1}^N \pm \varphi(-x_j) \)) such that \( \left\| \hat{\Phi}_\omega \right\|_{L^p} \sim N^{1/2} \), and conclude that the Hausdorff-Young inequality cannot hold when \( p > 2 \) by taking \( N \) sufficiently large.

Exercise 8. (★) Consider the circle \( T = \mathbb{R}/\mathbb{Z} \) and notice that Hölder’s inequality shows that when \( p \geq 2 \) we have \( \|f\|_{L^2(T)} \leq \|f\|_{L^p(T)} \) for all functions \( f \). Here we describe a class of functions for which the reverse inequality holds (which in particular implies that all the \( L^p \) norms are comparable).

Let \( f \) be a function on \( T \) such that \( \hat{f} \) is supported in the set \( \{3^k : k \in \mathbb{N}\} \); in particular, the Fourier series of \( f \) is given by \( \sum_{k \in \mathbb{N}} \hat{f}(3^k)e^{2\pi i 3^k x} \). You will show that for such functions one has
\[ \|f\|_{L^p(T)} \lesssim p^{1/2} \|f\|_{L^2(T)} \quad (3.9) \]
for any \( p \geq 2 \). The proof again rests on a randomisation trick enabled by Khintchine’s inequality.

i) Let \( (\epsilon_k(\omega))_{k \in \mathbb{N}} \) be as in the statement of Lemma 2.4. Assume that there exist Borel measures \( (\mu_\omega)_{\omega \in \Omega} \) such that \( \check{\mu}_\omega(3^k) = \epsilon_k(\omega) \) for any \( k \in \mathbb{N} \) and such that \( \|\mu_\omega\| \lesssim 1 \) uniformly in \( \omega \). Show that for functions \( f \) as above we can write
\[ f = \hat{f}_\omega \ast \mu_\omega, \]
where \( \hat{f}_\omega \) has Fourier series \( \sum_k \hat{f}(3^k)\epsilon_k(\omega)e^{2\pi i 3^k x} \).

ii) Show that, always under the assumption that the measures \( (\mu_\omega)_{\omega \in \Omega} \) exist, the above identity together with Young’s and Khintchine’s inequalities implies (3.9). The \( p^{1/2} \) constant comes from Exercise 5.

iii) It remains to show that the measures \( (\mu_\omega)_{\omega \in \Omega} \) really exist, and you will do so by constructing them explicitly. Let \( (p_k^\omega(\theta))_{k \in \mathbb{N}} \) be the collection of trigonometric polynomials given by
\[ p_k^\omega(\theta) := 1 + \frac{\epsilon_k(\omega)}{2} (e^{2\pi i 3^k \theta} + e^{-2\pi i 3^k \theta}) \]
and consider the limit \( \mu_\omega \) given by
\[ \mu_\omega(\theta) = \lim_{K \to \infty} \prod_{k=0}^K p_k^\omega(\theta). \]

Show that this limit exists in the weak sense, that is for any trigonometric polynomial \( q(\theta) \) the limit \( \lim_K (q, \prod_{k=0}^K p_k^\omega(\theta)) \) exists and is unique.

iv) Show that if an integer \( m \in \mathbb{Z} \) admits an expression of the form
\[ m = 3^{n_1} \pm \ldots \pm 3^{n_k} \]
for some integers \( 0 \leq n_1 < \ldots < n_k \) and some choice of signs, then such an expression is necessarily unique.

v) Show that, thanks to iv), \( \check{\mu}_\omega(3^k) = \epsilon_k(\omega) \), as desired.

vi) Show that \( \|\mu_\omega\| = \int |\mu_\omega| = 1 \) to conclude the proof.

vii) Let \( (N_k)_{k \in \mathbb{N}} \subset \mathbb{N} \) be a lacunary sequence, that is \( \liminf_{k \to \infty} N_{k+1}/N_k =: \rho > 1 \) (\( \rho \) is called the lacunarity constant of the sequence). Show that if \( \rho \geq 3 \) the proof above still works (step iv) in particular).

viii) If \( 3 > \rho > 1 \), show that one can decompose \( (N_k)_{k \in \mathbb{N}} \) into \( O((\log_3 \rho)^{-1}) \) sequences of lacunarity constant at least 3. Conclude that (3.9) holds for any function supported on a lacunary sequence, although the constant depends on the lacunarity constant of the sequence.
ix) Show that there exists a constant $C > 0$ such that the functions $f$ of the above kind satisfy the following interesting exponential integrability property:

$$\int_{\mathbb{T}} e^{C|f(\theta)|^2} \|f\|^2 \, d\theta \lesssim 1.$$  

[hint: Taylor-expand the exponential.]

**Exercise 9.** Show that if $\psi$ is chosen in such a way that

$$\sum_{j \in \mathbb{Z}} |\psi(2^{-j} \xi)|^2 = c$$

for all $\xi \neq 0$, then the converse of (2.3) is also true; that is, show that for any $1 < p < \infty$

$$\|f\|_{L^p} \lesssim_p \|\hat{\psi}f\|_{L^p}.$$ 

Finally, construct a function $\psi$ such that the above condition holds.

[hint: see the proof of Theorem 2.5]

**Exercise 10. (★)** Let $T_B$ denote the ball multiplier as in (2.6), that is the operator defined by $T_B f = \mathbf{1}_{B(0,1)} \hat{f}$. To keep things simple, we let the dimension be $d = 2$. Assume that $T_B$ is $L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ bounded for a certain $1 < p < \infty$ (in reality it is not, unless $p = 2$, but Fefferman’s proof of this fact proceeds by contradiction, so assume away). You will show that the following inequality would then also be true. Let $(v_j)_{j \in \mathbb{N}}$ be a collection of unit vectors in $\mathbb{R}$ and denote by $H_j$ the half-plane $\{x \in \mathbb{R}^2 : x \cdot v_j \geq 0\}$; finally, let $H_j$ be the operator given by $H_j f(\xi) = \mathbf{1}_{H_j}(\xi) \hat{f}(\xi)$ (you can see $H_j$ is essentially a Hilbert transform in direction $v_j$). If $T_B$ is $L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ bounded, then for any vector-valued function $(f_j)_{j \in \mathbb{N}} \in L^p(\mathbb{R}^2; \ell^2(\mathbb{N}))$ we have

$$\left\| \left( \sum_{j \in \mathbb{N}} |H_j f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^2)} \lesssim \left\| \left( \sum_{j \in \mathbb{N}} |f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^2)}.$$ 

The connection between $T_B$ and the $H_j$’s is that, morally speaking, a very large ball looks like a half-plane near its boundary.

i) Argue that the functions $f$ such that $\hat{f}$ is smooth and compactly supported are dense in $L^p$.

ii) Let $B_j^r$ denote the ball of radius $r$ and center $rv_j$ (thus $B_j^r$ is tangent to the boundary of $H_j$ at the origin for any $r > 0$) and let $T_j^r$ be the operator given by $T_j^r f = \mathbf{1}_{B_j^r} \hat{f}$. Show, using the Fourier inversion formula, that for functions $f$ as in i) it holds that

$$\lim_{r \to \infty} |H_j f(x) - T_j^r f(x)| = 0$$

for all $x$.

iii) Argue by Fatou’s lemma that

$$\left\| \left( \sum_{j \in \mathbb{N}} |H_j f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^2)} \leq \liminf_{r \to \infty} \left\| \left( \sum_{j \in \mathbb{N}} |T_j^r f_j|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^2)}$$

and conclude that, by i), it will suffice to bound the right-hand side uniformly in $r$.

iv) Let $T_{B_r}$ be the operator given by $T_{B_r} \hat{f} = \mathbf{1}_{B_r} \hat{f}$ (a rescaled ball multiplier); show that $T_{B_r}$ has the same $L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$-norm of $T_B = T_{B_1}$.

v) Show that $T_j^r = \text{Mod}_{-rv_j} T_{B_r} \text{Mod}_{rv_j}$, where $\text{Mod}_\xi f(x) := e^{-2\pi i \xi \cdot x} f(x)$.

vi) Combine v) with Exercise 8 to conclude.
**Exercise 11.** Let \( S \) be the Littlewood-Paley square function. Show that the inequality that would imply \( \|Sf\|_{L^p} \lesssim_p \|f\|_{L^p} \) by a duality argument is

\[
\left\| \sum_{j \in \mathbb{Z}} \Delta_j g_j \right\|_{L^{p'}} \lesssim_{p'} \left( \sum_{j \in \mathbb{Z}} |g_j|^2 \right)^{1/2}
\]

and not \( \|f\|_{L^{p'}} \lesssim_p \|Sf\|_{L^{p'}} \) as one could naively expect.

**Exercise 12.** Show, using both \( \lesssim_p \) and \( \gtrsim_p \) inequalities of \([1]\), that the operators

\[
T_\omega f := \sum_{j \in \mathbb{Z}} \epsilon_j(\omega) \Delta_j f
\]

are \( L^p(\mathbb{R}) \to L^p(\mathbb{R}) \) bounded for \( 1 < p < \infty \), with constant independent of \( \omega \). Notice that, unlike the operators \( \tilde{T}_\omega \), the operators \( T_\omega \) are not Calderón-Zygmund operators in general (prove this for some special choice of signs), and therefore we really need to use Littlewood-Paley theory to prove they are bounded.

**Exercise 13.** Assume that function \( m(\xi) \) satisfies \( \|m\|_{L^\infty} < \infty \) and that it has bounded variation over all of \( \mathbb{R} \) (that is \( \int_{\mathbb{R}} |d\nu(\xi)| < \infty \)). Show, without appealing to the Marcinkieicuwicz multiplier theorem, that \( m \) defines a multiplier which is \( L^p(\mathbb{R}) \to L^p(\mathbb{R}) \) bounded for all \( 1 < p < \infty \).

[Hint: simplify the proof of Theorem 3.2 as much as you can.]

**Exercise 14.** Show that the pointwise condition ii') in Remark 2 implies condition ii) in Theorem 3.1.

**Exercise 15.** (★) In this exercise you will prove Theorem 3.1 in dimensions \( d > 1 \). Actually, the notation required becomes nightmarish pretty quickly, and thus we will content ourselves with proving the case \( d = 2 \), since it already contains the full generality of the argument. Let \( T = T_m \) and \( m \) satisfy the conditions stated in the theorem. The proof is essentially a repetition of the argument given for Theorem 3.2.

i) Show that, by Theorem 2.7, it will suffice to show

\[
\left\| \left( \sum_{k \in \mathbb{Z}^2} |\Delta_{R_{k_1} k_2} Tf|^2 \right)^{1/2} \right\|_{L^p} \lesssim_p \left( \sum_{k \in \mathbb{Z}^2} |\Delta_{R_{k_1} k_2} f|^2 \right)^{1/2}
\]

under the assumption that \( \hat{f} \) is supported in the first quadrant \([0, \infty) \times [0, \infty)\).

ii) Show that in each rectangle \( R_{k_1, k_2} \) we can write

\[
m(\xi, \eta) = m(2^{k_1}, 2^{k_2}) + \int_{2^{k_1}}^\xi \partial_{\xi} m(\zeta_1, 2^{k_2}) d\zeta_1 + \int_{2^{k_2}}^\eta \partial_{\eta} m(2^{k_1}, \zeta_2) d\zeta_2 + \int_{2^{k_1}}^\xi \int_{2^{k_2}}^\eta \partial_{\xi} \partial_{\eta} m(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2.
\]

iii) Show that by ii) and Fubini we have

\[
\Delta_{R_{k_1} k_2} Tf = m(2^{k_1}, 2^{k_2}) \Delta_{R_{k_1} k_2} f + \int_{2^{k_1}}^{2^{k_1+1}} \partial_{\xi} m(\zeta_1, 2^{k_2}) \Delta_{\zeta_1}^{(1)} f d\zeta_1
\]

\[
+ \int_{2^{k_2}}^{2^{k_2+1}} \partial_{\eta} m(2^{k_1}, \zeta_2) \Delta_{\zeta_2}^{(2)} f d\zeta_2
\]

\[
+ \int_{2^{k_1}}^{2^{k_1+1}} \int_{2^{k_2}}^{2^{k_2+1}} \partial_{\xi} \partial_{\eta} m(\zeta_1, \zeta_2) \Delta_{\zeta_1}^{(1)} \Delta_{\zeta_2}^{(2)} f d\zeta_1 d\zeta_2.
\]

iv) Use condition i) of the theorem to dispense with the first term at the right-hand side of the last expression.
Exercise 17. Show that the range

\[ C \left( \int_{2^k}^{2^{k+1}} \left| \Delta_{[\xi_1, 2^{k+1}]} f \right|^2 |\partial \xi m(\xi_1, 2^{k+1})| d\xi_1 \right) + \int_{2^{k+1}}^{2^{k+2}} \left| \Delta_{[\xi_2, 2^{k+2}]} f \right|^2 |\partial \eta m(2^{k+1}, \xi_2)| d\xi_2 + \int_{2^{k+1}}^{2^{k+2}} \int_{2^{k+1}} \left| \Delta_{[\xi_3, 2^{k+1}]} x \cdot [\xi_2, 2^{k+2}]} f \right|^2 |\partial \mu m(\xi_1, \xi_2, \xi_3)| d\xi_1 d\xi_2 \].

v) Use condition ii) and Cauchy-Schwarz to show that the (square of) the remaining terms is bounded by

\[ C \left( \int_{2^k}^{2^{k+1}} \left| \Delta_{[\xi_1, 2^{k+1}]} f \right|^2 |\partial \xi m(\xi_1, 2^{k+1})| d\xi_1 \right) + \int_{2^{k+1}}^{2^{k+2}} \left| \Delta_{[\xi_2, 2^{k+2}]} f \right|^2 |\partial \eta m(2^{k+1}, \xi_2)| d\xi_2 + \int_{2^{k+1}}^{2^{k+2}} \int_{2^{k+1}} \left| \Delta_{[\xi_3, 2^{k+1}]} x \cdot [\xi_2, 2^{k+2}]} f \right|^2 |\partial \mu m(\xi_1, \xi_2, \xi_3)| d\xi_1 d\xi_2 \].

vi) Find measure spaces \((\Gamma_j, d\mu_j)\) and collections of rectangles or intervals for \( j = 1, 2, 3 \) such that each of the 3 terms above can be treated by Proposition 2.1 as in the proof of the \( d = 1 \) case.

vii) Conclude using condition ii) one last time.

Exercise 16. Show that the multiplier given by

\[ m(\xi_1, \ldots, \xi_d) := \frac{|\xi_1|^{\alpha_1} \cdots |\xi_d|^{\alpha_d}}{(\xi_1^2 + \cdots + \xi_d^2)^{\alpha_2}} \]

where \( \alpha_j > 0 \) and \( |\alpha| := \alpha_1 + \cdots + \alpha_d \) is a Marcinkiewicz multiplier.

Exercise 17. Show that the range \( \frac{\alpha_1}{d+1} < p \leq \infty \) for the boundedness of the spherical maximal function \( \mathcal{M}_{S^{d-1}} \) is sharp, in the sense that \( \mathcal{M}_{S^{d-1}} \) is not bounded on \( L^p(\mathbb{R}^d) \) for any \( 1 \leq p < \frac{\alpha_1}{d+1} \) (find a counterexample).

Exercise 18. Let the dimension be \( d = 3 \) and let \( A_r \) denote the spherical average \( A_r f(x) := \int_{S^2} f(x - ty) d\sigma(y) \), where \( d\sigma \) is the normalised surface measure on \( S^2 \). Show that \( u(x, t) := A_t g(x) \) solves the wave equation

\[ \begin{cases} \partial^2_t u - \Delta u = 0, \\ u(x, 0) = 0, \\ \partial_t u(x, 0) = g(x), \end{cases} \]

with \( x \in \mathbb{R}^3, t \geq 0 \); here we assume \( g \in C^2(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \). This is an instance of Huygens’ principle, and more in general if the initial data becomes \( u(x, 0) = f(x) \) with \( f \in C^2 \), the complete solution is \( t A_t g + \frac{d}{d t}(t A_t f) \). Similar formulas hold in higher dimensions, but only when \( d \) is odd they involve only the spherical averages \( A_r \). When \( d \) is even, one needs to average over balls instead (that is, the classical Huygens’ principle fails in even dimensions).

i) Using Stokes theorem, show that

\[ \frac{d}{d t} A_t g(x) = \frac{1}{t^2} \Delta \left( \int_{|y| \leq t} g(x - y) dy \right). \]

ii) Using polar coordinates, show that

\[ \int_{|y| \leq t} g(x - y) dy = \int_0^t s^2 A_s g(x) ds. \]

iii) Use i)-ii) to show that

\[ \frac{d}{d t} \left( t^2 \frac{d}{d t} A_t g(x) \right) = t^2 \Delta A_t g(x). \]

iv) Show that for any \( F \) twice differentiable it holds that \( \frac{1}{t} \frac{d}{d t} \left( t^2 \frac{d}{d t} F(t) \right) = \left( \frac{d}{d t} \right)^2 (t F(t)) \), and combine this with iii) to show that the wave equation is satisfied.
v) It remains to check the initial conditions. Using Theorem 3.3 argue that
\[ \lim_{t \to 0} A_r g(x) = g(x) \text{ for every } x, \] and then use this fact together with i) to argue that the initial conditions are indeed satisfied in the limit \( t \to 0 \) (and \( u, \partial u \) are continuous in \( t \geq 0 \)).

Exercise 19. Let \( \psi \) be a smooth function compactly supported in the ball \( B(0,1) \) of \( \mathbb{R}^d-1 \). In this exercise you will show that
\[
\left| \int_{\mathbb{R}^{d-1}} e^{i(x',x_d) \cdot (\xi,|\xi|^2)} \psi(\xi) d\xi \right| \lesssim (1 + |x'| + |x_d|)^{-(d-1)/2} \tag{3.10}
\]
for any \((x',x_d) \in \mathbb{R}^{d-1} \times \mathbb{R} = \mathbb{R}^d\). It is a matter of doing a smooth partition of unity on \( \mathbb{R}^{d-1} \) and a change of variables with a suitable diffeomorphism to show that the above implies \( |\sigma(\xi)| + |\nabla \sigma(\xi)| \lesssim (1 + |\xi|)^{-(d-1)/2} \); however, we will content ourselves with the object above (you can see it as the Fourier transform of a measure supported on the elliptic paraboloid parametrised by \((\xi,|\xi|^2)\), instead of the sphere; but the two surfaces are locally the same).

1. Square the left-hand side of (3.10) and show by a change of variables that the result equals
\[
I = \int_{[\mathbb{R}^{d-1}] 	imes \mathbb{R}^{d-1}} e^{i(x',x_d) \cdot (\eta_1,\eta_2)} \Psi(\eta_1,\eta_2) d\eta_1 d\eta_2
\]
for some smooth function \( \Psi \) compactly supported in \( B(0,2) \times B(0,2) \).

2. You will show (3.10) in two complementary cases. First let \( |x'| < 200|x_d| \). Show that
\[
|I| \leq \int_{[\mathbb{R}^{d-1}]} |\mathcal{F} \Psi(\eta_1, x_d \eta_1) d\eta_1,\]
where \( \mathcal{F} \) denotes the Fourier transform in the second variable. Argue that
\[
|\mathcal{F}(\xi)\Psi(\eta_1, x_d \eta_1)| \lesssim_N (1 + |\eta_1| + |x_d| |\eta_1|)^{-N}
\]
for any arbitrarily large \( N > 0 \) and conclude that this shows \( |I| \lesssim (1 + |x_d|)^{-(d-1)} \sim (1 + |x'| + |x_d|)^{-(d-1)}, \) thus proving (3.10) in this regime.

3. Let now \( |x'| \geq 200|x_d| \). Show that
\[
|I| \leq \int_{[\mathbb{R}^{d-1}]} |\mathcal{F} \Psi(x', x_d \eta_2, \eta_2) d\eta_2|,
\]
next argue that
\[
|\mathcal{F}(\xi)\Psi(x', x_d \eta_2, \eta_2)| \lesssim_N (1 + |x'| + |x_d| |\eta_2|)^{-N} \lesssim_N (1 + |x'|)^{-N}
\]
for any \( N > 0 \) and show that this proves (3.10) in this case as well (even with arbitrarily large exponents).

Exercise 20. Let \( \psi_j \) be as in the proof of Theorem 3.3 Show that, thanks to the fact that \( |\hat{\psi}(x)| \lesssim_N (1 + |x|)^{-N} \) for any \( N > 0 \), one has the rough bound
\[
|A_r \hat{\psi}_j(x)| \lesssim 2^j (1 + |x|)^{-10d}.
\]
for any \( r \sim 1 \) and any \( x \in \mathbb{R}^d \).

Exercise 21. (★) In this exercise you will finish the proof of Theorem 3.3 Here you will need to use a smooth square function at some point, but other than that the proof is just a repetition of the one given for \( M_{L^d-1} \).

i) We see \( M_{L^d-1} \) as
\[
\sup_{k \in \mathbb{Z}} \sup_{2^k \leq r \leq 2^{k+1}} |A_r f|.
\]
For a fixed \( k \in \mathbb{Z} \) let \( P_{\leq k} = \sum_{j \leq k} P_j \) be the operator that will take on the rôle of \( P_{low} \), so that
\[
f = P_{\leq k} f + \sum_{j \in \mathbb{N}} P_{j+k} f.
\]
Show that it suffices to prove for any $d/(d-1) < p < 2$

$$
\| \sup_{k \in \mathbb{Z}} \sup_{2^{-k} \leq r \leq 2^{-k+1}} |A_r P_{\leq k} f| \|_{L^p} \lesssim_p \| f \|_{L^p}, \quad (3.11)
$$

$$
\| \sup_{k \in \mathbb{Z}} \sup_{2^{-k} \leq r \leq 2^{-k+1}} |A_r P_{j+k} f| \|_{L^p} \lesssim_p 2^{-\delta j} \| f \|_{L^p}, \quad (3.12)
$$

for some $\delta = \delta_{p,d} > 0$.

ii) Show that for $r \sim 2^{-k}$ one has

$$
|A_r P_{\leq k} f| \lesssim M f
$$

uniformly in $r, k$ and conclude [3.11]. You can simply rescale the argument given for the analogous part of $M_{\geq -1}$, that is for $k = 0$.

iii) Show that for $r \sim 2^{-k}$ and $j > 0$ one has

$$
|A_r P_{j+k} f| \lesssim 2^j M f
$$

uniformly in $r$ and $k$ (again, just rescale the argument used when $k = 0$). Conclude that

$$
\| \sup_{k \in \mathbb{Z}} \sup_{2^{-k} \leq r \leq 2^{-k+1}} |A_r P_{j+k} f| \|_{L^q} \lesssim_p 2^j \| f \|_{L^q} \quad (3.13)
$$

for any $q > 1$.

iv) Show, using the decay of $d\sigma$, that when $r \sim 2^{-k}$

$$
\|A_r P_{j+k} f\|_{L^2} \lesssim 2^{-j(d-1)/2} \| f \|_{L^2}.
$$

v) Show, using the decay of $\nabla d\sigma$, that when $r \sim 2^{-k}$

$$
\left\| \frac{d}{dr} A_r P_{j+k} f \right\|_{L^2} \lesssim 2^k 2^{-j(d-3)/2} \| f \|_{L^2}.
$$

vi) Show, using a suitably rescaled inequality [3.7], that for any $k \in \mathbb{Z}$

$$
\| \sup_{2^{-k} \leq r \leq 2^{-k+1}} |A_r P_{j+k} f| \|_{L^2} \lesssim 2^{-j(d-2)/2} \| f \|_{L^2}. \quad (3.14)
$$

vii) Let $\tilde{\psi}$ be a smooth function compactly supported in the annulus $\{ \xi \in \mathbb{R}^d : 1/4 \leq |\xi| \leq 4 \}$ and identically equal to 1 in the annulus $\{ \xi \in \mathbb{R}^d : 1/2 \leq |\xi| \leq 2 \}$; finally, let $\tilde{P}_j$ be the annular projection given by $\tilde{P}_j f(\xi) = \tilde{\psi}(2^{-j}\xi) \tilde{f}(\xi)$ (thus they are just a slight modification of the projections in Section 2.4). Show that in (3.14) we can replace $\| f \|_{L^2}$ in the right-hand side by $\| \tilde{P}_{j+k} f \|_{L^2}$.

viii) Replace the supremum in $k \in \mathbb{Z}$ by the $L^2(\mathbb{Z})$ sum and show that, by Proposition 2.6, we have

$$
\| \sup_{k \in \mathbb{Z}} \sup_{2^{-k} \leq r \leq 2^{-k+1}} |A_r P_{j+k} f| \|_{L^2} \lesssim 2^{-j(d/2-1)} \| f \|_{L^2}. \quad (3.15)
$$

ix) Interpolate between (3.13) and (3.15) to show that (3.12) holds. This concludes the proof.

**Exercise 22. (★)** The Carleson operator for $\mathbb{R}$ (already mentioned in previous lectures for $\mathbb{T}$) is given by

$$
\mathcal{C} f(x) := \sup_{N>0} \left| \int_{[-N,N]}(\xi) \tilde{f}(\xi) e^{2\pi i \xi x} d\xi \right|.
$$

Recall that its $L^p \to L^{p,\infty}$ boundedness for a certain $p$ implies the a.e. convergence of the integral above to $f(x)$ as $N \to \infty$ when $f \in L^p$. The methods developed in these notes are not powerful enough to show the boundedness of $\mathcal{C}$, but they are
enough for a simplified version of it. Indeed, in this exercise you will prove the $L^p$ boundedness of the lacunary Carleson operator

$$C_{\text{lac}}f(x) := \sup_{k \in \mathbb{Z}} \left| \int_{1 - 2^k, 2^k} (\xi) \hat{f}(\xi)e^{2\pi i \xi x} d\xi \right|.$$  

Fix $p$ such that $1 < p < \infty$.

1. First, consider the additional simplification given by replacing the characteristic function $1_{[-N,N]}(\xi)$ by the smooth function $\varphi(\xi/N)$ where $\varphi$ is a non-negative smooth function compactly supported in $[-1,1]$ and identically equal to 1 on $[-1/2, 1/2]$. Show that the associated maximal operator

$$\hat{C}f(x) := \sup_{N > 0} \left| \int \varphi(\xi/N) \hat{f}(\xi)e^{2\pi i \xi x} d\xi \right|$$

is bounded pointwise by a constant multiple of the Hardy-Littlewood maximal function $Mf(x)$ and conclude $L^p \to L^p$ boundedness of $\hat{C}$.

2. Show that it suffices to prove the $L^p$ boundedness of the operator

$$\Theta_{\text{lac}}f(x) := \sup_{k \in \mathbb{Z}} \left| \int (1_{[-2^k,2^k]}(\xi) - \varphi(2^{-k}\xi)) \hat{f}(\xi)e^{2\pi i \xi x} d\xi \right|$$

to conclude the same for $\Theta_{\text{lac}}$.

3. Show that there is a smooth non-negative function $\theta$ compactly supported in $[1/2, 2] \cup [-2, -1/2]$ and such that

$$1_{[-2^k,2^k]}(\xi) - \varphi(2^{-k}\xi) = 1_{[2^k-1,2^k]}(\xi) - 2^{k-1}(\xi) \cdot \theta(2^{-k}\xi).$$

4. Let $\Theta_k$ denote the frequency projections associated to $\theta$, that is $\Theta_k f(\xi) = \theta(2^{-k}\xi) \hat{f}(\xi).$ Show that it suffices to show that the square function

$$\mathcal{S}f := \left( \sum_{k \in \mathbb{Z}} |\Delta_k \Theta_k f|^2 \right)^{1/2}$$

is $L^p \to L^p$ bounded to conclude the same for $\Theta_{\text{lac}}$.

5. Show that $\mathcal{S}$ is indeed $L^p \to L^p$ bounded.

6. Argue that the argument generalises to treat the operators

$$C'_{\text{lac}}f(x) := \sup_{k \in \mathbb{Z}} \left| \int 1_{[-N_k,N_k]}(\xi) \hat{f}(\xi)e^{2\pi i \xi x} d\xi \right|$$

where $(N_k)_{k \in \mathbb{N}}$ is any fixed lacunary sequence - that is, a sequence that satisfies $\lim \inf_{k \to \infty} N_{k+1}/N_k =: \rho > 1$. However, the constant produced by the proof will end up depending on $\rho$.

**Exercise 23.** ($\star$) Recall that in the lecture on Calderón-Zygmund theory you have seen how, thanks to the boundedness of the Riesz transforms (defined by $\overline{R}_j f(\xi) = (\xi_j/|\xi|) \hat{f}(\xi)$), we can control the $L^p$ norms of all the partial derivatives of order 2 by the Laplacian alone: for any $i,j \in \{1, \ldots, d\}$

$$\|\partial_x \partial_x u\|_{L^p} \lesssim_p \|\Delta u\|_{L^p} \quad \text{for all } 1 < p < \infty.$$  

With a little spherical harmonics theory, this can be generalised to generic homogeneous elliptic differential operators, always within the framework of Calderón-Zygmund theory. More precisely, let $Q(X) \in \mathbb{C}[X_1, \ldots, X_d]$ be a homogeneous elliptic polynomial of degree $k$, that is

1) $Q(\lambda X) = \lambda^k Q(X)$ for all $\lambda \neq 0$;
2) the zero set $Z(Q)$ of $Q(X)$ consists only of the origin.
If \( Q(X) = \sum_{\alpha \in \mathbb{N}^d : |\alpha| = k} c_\alpha X^\alpha \), the homogeneous elliptic differential operator \( Q(\partial) \) is defined as
\[
Q(\partial) := \sum_{\alpha \in \mathbb{N}^d : |\alpha| = k} c_\alpha \partial^\alpha.
\]
Then one can show that for all multi-indices \( \alpha \in \mathbb{N}^d \) such that \( |\alpha| = k \) and (at least) for all the functions \( u \) of class \( C^k \) and compact support, one has
\[
\| \partial^\alpha u \|_{L^p} \lesssim_{p,Q} \| Q(\partial) u \|_{L^p} \quad \text{for all } 1 < p < \infty.
\]
The point is that the operator \( T \) such that \( \partial^\alpha = T \circ Q(\partial) \) exists and is a bounded Calderón-Zygmund singular integral operator (or a composition of such operators).

Now, as a further application of Littlewood-Paley theory, you will show that the result above generalises even further. Indeed, you will show that if \( P(X) \in \mathbb{C}[X_1, \ldots, X_d] \) is a polynomial of degree \( k \) whose top-degree component is elliptic, then for all multi-indices \( \alpha \in \mathbb{N}^d \) such that \( |\alpha| \leq k \) and (at least) for all the functions \( u \) of class \( C^k \) and compact support, one has
\[
\| \partial^\alpha u \|_{L^p} \lesssim_{p,P} \| P(\partial) u \|_{L^p} + \| u \|_{L^p} \quad \text{for all } 1 < p < \infty.
\]

i) Let \( Z(P) \) denote the zero set of \( P \). Show that \( Z(P) \) is contained in a ball \( B(0,R) \) for some large \( R \) depending on \( P \).

ii) Let \( \varphi \) be a smooth function compactly supported in a neighbourhood of \( Z(P) \) and that vanishes outside \( B(0,2R) \).

iii) With \( \xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d} \), show that \( m_1(\xi) := \xi^\alpha \varphi(\xi) \) is a Schwartz function and deduce that \( T_{m_1} \) is \( L^p \to L^p \) bounded for \( 1 < p < \infty \).

iv) Show that \( m_2(\xi) := (1 - \varphi(\xi)) P(\xi) \hat{\varphi}(\xi) \) is a Marcinkiewicz multiplier and therefore \( T_{m_2} \) is \( L^p \to L^p \) bounded for \( 1 < p < \infty \).

v) Decompose \( \xi^\alpha \hat{f}(\xi) = m_1(\xi) \hat{f}(\xi) + m_2(\xi) P(\xi) \hat{f}(\xi) \) and use this to conclude the inequality.