SOME NOTES ON RESTRICTION THEORY

KONSTANTIN MERZ

Abstract. In these notes, we review the state of progress on the restriction problem in harmonic analysis with an emphasis on the developments of the past decade or so on the euclidean space version of these problems for spheres and other hypersurfaces. As the field is quite large, we will merely give the main ideas and developments in this area.

The restriction problem is connected to many other conjectures, most notably the Kakeya and Bochner–Riesz conjectures, as well as PDE conjectures such as the local smoothing conjecture which will be discussed as well.

These notes are mostly based on Tao's famous review [185], his lecture notes on restriction problems [179], the lecture notes by Wolff on harmonic analysis [200], recent lecture notes by Hickman and Vitturi on decoupling theory [107], and the introductory review [171] by Stovall.

Contents

Date: August 10, 2023.

1. The restriction problem – some background

From now on, we fix $d \geq 2$ and remark that all constants A or a are allowed to depend on d (although it would be interesting to track the precise dependence of the constants on the dimension as $d \to \infty$). The Fourier transform of a function f on $\mathbb{R}^d$ is formally defined as

$$\hat{f}(\xi) := \int_{\mathbb{R}^d} f(x) \mathrm{e}^{-2\pi i x \cdot \xi} \, dx \, .$$

By the Riemann–Lebesgue lemma we know that $\hat{f}$ is a continuous bounded function on $\mathbb{R}^d$ which vanishes at infinity if $f \in L^1(\mathbb{R}^d)$. In particular, $\hat{f}$ can be meaningfully restricted to any subset S of $\mathbb{R}^d$, thereby creating a continuous bounded function on S.

For applications, the above definition needs to be extended to a larger class of functions. For $f \in L^1 \cap L^2$, the Plancherel theorem states that $\|f\|_2 = \|\hat{f}\|_2$ and since $L^1 \cap L^2$ is dense in L^2, the Fourier transform extends uniquely to a bounded linear operator of L^2 onto itself. By interpolation, we obtain the Hausdorff–Young inequality which states that for $1 \leq p \leq 2$, this extension maps L^p boundedly into $L^{p'}$ and obeys $\|\hat{f}\|_{p'} \leq A_{p,d} \|f\|_p$. This range of $L^p \to L^q$

estimates is the best possible; for the sharp constant $A_{p,d}$, see Beckner [4] and for extremizers, see Lieb [131].

For $f \in L^p$, $p > 1$, $\hat{f}$ is usually interpreted as an $L^{p'}$ limit, $\hat{f} = \lim_{n\to\infty} \hat{f}_n$ where f_n is a sequence of integrable functions converging to f in L^p. By the Hausdorff–Young inequality, one can therefore restrict $\hat{f}$ to any set S of *positive* measure. However, the above interpretation leads to an obvious obstruction to restricting a Fourier transform to *sets of Lebesgue measure zero.* Indeed $L^{p'}$ consists of equivalence classes within which its members are allowed to differ off of sets of measure zero, i.e., it makes no sense to define Fourier restriction to a set of measure zero as a simple composition. In particular, there is no meaningful way to restrict L^2 functions to any set S of measure zero.

In 1967 Stein made the surprising discovery (unpublished work) that when such sets contain "sufficient curvature" (see also Subsection 3.4), then one can indeed restrict the Fourier transform of L^p functions for certain $p > 1$. This lead to the *restriction problem* [163]: for which sets $S \subseteq \mathbb{R}^d$ and which $1 \leq p \leq q \leq \infty$ can the Fourier transform of an L^p function be meaningfully restricted, i.e.,

$$\|\hat{f}|_S\|_{L^q(S)} \leq A_{p,q,d}\|f\|_{L^p(\mathbb{R}^d)}$$

for smooth, compactly supported f?

Of course, there are infinitely many such sets to consider, but we will focus on sets S which are hypersurfaces, or compact subsets of hypersurfaces. In particular, we shall be interested in the *sphere*

$$S_{\text{sphere}} := \{\xi \in \mathbb{R}^d : |\xi| = 1\}\,,$$

the *paraboloid*

$$S_{\text{parab}} := \{\xi \in \mathbb{R}^d : \xi_d = |\underline{\xi}|^2/2\}\,,$$

and the *cone*

$$S_{\text{cone}} := \{\xi \in \mathbb{R}^d : \xi_d = |\underline{\xi}|\}$$

where $\xi = (\underline{\xi}, \xi_d) \in \mathbb{R}^{d-1} \times \mathbb{R} \equiv \mathbb{R}^d$. These three surfaces are model examples of hypersurfaces with curvature, though of course the cone differs from the sphere and the paraboloid in that it has one vanishing principal curvature. These three surfaces also enjoy a large group of symmetries (the orthogonal group, the parabolic scaling and Galilean groups, and the Poincaré group, respectively). Moreover, these hypersurfaces are intimately related (via the Fourier transform) to certain PDEs, namely the Helmholtz equation, the Schrödinger equation, and the wave equation, respectively. This is going to be the topic of Section 13.

Organization. The rest of the notes is structured as follows. In the next section, we will use a duality argument to reformulate the restriction problem as an "extension problem" which (as of this writing) is a more convenient point of view to think of the problem. In Section 3, we will find two necessary conditions for the restriction problem which lead to the *restriction conjecture.* In Sections 6 and 7 we will describe two classical tools to tackle the restriction conjecture. A more recent approach via Littlewood–Paley theory is discussed in Section 8. Connected to Littlewood–Paley theory is a recently established tool by Bourgain and Demeter [29, 30] (see also the references in [107]), namely ℓ^2 decoupling, which is the topic of Section 21. In Sections 13, 15, and 16 we will illuminate certain relations between the restriction problem and conjectures concerning nonlinear, dispersive PDEs, the Kakeya, and the Bochner–Riesz conjectures.

2. Restriction and extension estimates

From now on let S be a compact subset (but with non-empty interior) of one of the above surfaces S_{sphere}, S_{parab}, or S_{cone}. We endow S with a canonical measure $d\sigma$. For the sphere, it is the surface measure; for the paraboloid, it is the pullback of the $d-1$-dimensional Lebesgue measure $d\underline{\xi}$ under the projection map $\xi \mapsto \underline{\xi}$; for the cone it is the pullback of $d\underline{\xi}/|\xi|$ as it is Lorentz invariant.

In order to restrict $\hat{f}$ to S, it will suffice to prove an *a priori* "restriction estimate" of the form

$$\|\hat{f}|_S\|_{L^q(S,d\sigma)} \leq A_{p,q,S}\|f\|_{L^p(\mathbb{R}^d)} \tag{2.1}$$

for all C_c^∞ or Schwartz functions f and some $1 \leq q \leq \infty$, since one can then use density arguments to obtain a continuous restriction operator from $L^p(\mathbb{R}^d)$ to $L^q(S, d\sigma)$ which extends the *restriction operator* $\mathcal{R} : f \mapsto \hat{f}|_S$ for such nice functions. (Finding the sharp value of $A_{p,q,S}$ in (2.1) is another interesting difficult problem, which has only been solved in a few cases so far).

We will denote by $R_S(p \to q)$ the statement that (2.1) holds for all f. From the introductory remarks on the Hausdorff–Young inequality (the *faster* a function decays, i.e., f lives in *low* L^p spaces, the *smoother* is its Fourier transform, i.e., $\hat{f}$ lives in *high* L^q spaces), we see that $R_S(1 \to q)$ holds for all $1 \leq q \leq \infty$ by Hölder's inequality while $R_S(2 \to q)$ fails for all $1 \leq q \leq \infty$. The interesting question is what happens for intermediate values of p, i.e., our aim is to find the highest value of p (the slowest decay of f) and q (greatest smoothness of $\hat{f}$) such that the restriction estimate (2.1) still holds. (Observe the implication $R_S(p \to q) \Rightarrow R_S(\tilde{p} \to \tilde{q})$ for all $\tilde{p} \leq p$ and $\tilde{q} \leq q$ by Sobolev and Hölder inequalities.)

The dual of the restriction operator $\mathcal{R}_S$ is the *extension operator*

$$(\mathcal{R}_S)^*F(x) =: \mathcal{E}_S F(x) = (F\,d\sigma)^\vee = \int_S F(\xi)e^{2\pi i x\cdot\xi}d\sigma(\xi)\,.$$

A simple duality argument based on Parseval's identity (think of a change of variables to understand $(Fd\sigma)^\vee$ better, too)

$$\begin{aligned}
\sup_{\|f\|_{L^p(\mathbb{R}^d)}=1}\|\hat{f}|_S\|_{L^q(S,d\sigma)} &= \sup_{\|f\|_{L^p(\mathbb{R}^d)}=1}\ \sup_{\|F\|_{L^{q'}(S,d\sigma)}=1}\int_S F(\xi)\overline{\hat{f}(\xi)}\,d\sigma(\xi)\\
&= \sup_{\|F\|_{L^{q'}(S,d\sigma)}=1}\ \sup_{\|f\|_{L^p(\mathbb{R}^d)}=1}\int_{\mathbb{R}^d}(Fd\sigma)^\vee(x)\overline{f(x)}\,dx = \sup_{\|F\|_{L^{q'}(S,d\sigma)}=1}\|(Fd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)}
\end{aligned} \tag{2.2}$$

shows that the restriction estimate (2.1) is equivalent to the following *extension estimate*

$$\|(Fd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)} \leq A_{p,q,d}\|F\|_{L^{q'}(S,d\sigma)} \tag{2.3}$$

for all smooth functions F on S. We use $R_S^*(q' \to p')$ to denote the statement that the estimate (2.3) holds. Due to the smoothness of F one may use stationary phase arguments to obtain asymptotics for $(Fd\sigma)^\vee$, see also [167, Chapter VIII, Proposition 6]. However, such asymptotics depend on the smooth norms of F, not just the $L^{q'}(S)$ norm, and so do not imply estimates of the form (2.3). In this sense, one can think of extension estimates as a more general way to control oscillatory integrals since only magnitude bounds on $F(\xi)$ and no bounds on derivatives are required.

Understanding the extension operator better. Let us clarify at this stage the meaning of the restriction and extension operators. Suppose $f \in L^p(\mathbb{R}^d)$ and $F \in L^{q'}(S, d\sigma)$ as in the above duality argument. After rotating and translating S in the ambient space $\mathbb{R}^d$, we may assume (since S is compact) that S is given as the graph

$$\xi_d = \varphi(\xi_1, ..., \xi_{d-1})$$

where $\varphi \in C_c^\infty(\mathbb{R}^{d-1})$. This allows us to write the measure as

$$d\sigma(\xi) = (1 + |\nabla\varphi|^2)^{1/2}d\xi_1, ..., d\xi_{d-1}\,,$$

which is called the *euclidean (or induced) surface measure.* (Note that S is the level set of a function $\Psi : \mathbb{R}^d \to \mathbb{R}$ and that the measure is actually given by

$$d\sigma(\xi) = \frac{|(\nabla\Psi)(\xi_1, ..., \xi_d)|}{|\partial\Psi/\partial\xi_d|} d\xi_1, ..., d\xi_{d-1}.$$

(Compare this to the "canonical measure" $d\Sigma_\lambda(\xi) = |\nabla\Psi(\xi)|^{-1} d\sigma(\xi)$ which equals, locally at least, $|\partial\Psi/\partial\xi_d|^{-1}\, d\xi'$ in Yafaev [202, Chapter 2, Formula (1.4) and p. 111].) Using $\xi_d = \varphi(\xi_1, ..., \xi_{d-1})$ and the chain rule (remember $(\partial\xi_d/\partial\xi_i)_{i=1}^{d-1} = \nabla\varphi$), we (formally) have

$$\left(\frac{|(\nabla\Psi)(\xi_1, ..., \xi_d)|}{|\partial\Psi/\partial\xi_d|}\right)^2 = \sum_{i=1}^{d} \frac{|\partial\Psi/\partial\xi_i|^2}{|\partial\Psi/\partial\xi_d|^2} = 1 + \sum_{i=1}^{d-1} \left|\frac{\partial\Psi/\partial\xi_i}{\partial\Psi/\partial\xi_d}\right|^2 = 1 + \sum_{i=1}^{d-1} \left|\frac{\partial\xi_d}{\partial\xi_i}\right|^2 = 1 + |\nabla\varphi|^2$$

which yields the previous representation. Alternatively, using the implicit function theorem, we know that locally $\xi_d = \varphi(\xi')$, whenever $(\xi', \xi_d) \in S$ where S denoted the level set of $\Psi : \mathbb{R}^d \to \mathbb{R}$. Thus, locally, $\Psi(\xi) = 0 = \xi_d - \varphi(\xi') = 0$, i.e., $\nabla\Psi(\xi) = (-\nabla\varphi(\xi'), 1)$ and $\partial\Psi/\partial\xi_d = 1$ on S. Therefore, $|\nabla\Psi(\xi)|/|\partial\Psi/\partial\xi_d| = (1 + |\nabla\varphi(\xi')|^2)^{1/2}$ and in particular $d\sigma(\xi) = (1 + |\nabla\varphi(\xi')|^2)^{1/2} d\xi'$.)

Now, using the above representation, abbreviating $\psi(\xi) = \sqrt{1 + |\nabla\varphi(\xi)|^2}$ and $\tilde{\xi} \equiv (\xi, \xi_d) = (\xi_1, ..., \xi_{d-1}, \xi_d) \in \mathbb{R}^d$, we may write the left side of (2.2) as

$$\int_S \hat{f}(\xi) F(\xi)\, d\sigma(\xi) = \int_{\mathbb{R}^{d-1}} \hat{f}(\xi) F(\xi) \psi(\xi)\, d\xi = \int_{\mathbb{R}^d} \hat{f}(\tilde{\xi}) F(\tilde{\xi}) \psi(\tilde{\xi}) \mathbf{1}_{\xi_d = \varphi(\xi)}\, d\tilde{\xi}$$

where $\mathbf{1}_{\xi_d = \varphi(\xi)}$ is to be understood as the one dimensional Dirac delta function which forces $\xi_d = \varphi(\xi)$. Using Parseval's theorem (in $L^2(\mathbb{R}^d)$), the right side of the last formula equals (with $x \in \mathbb{R}^d$)

$$\int_{\mathbb{R}^d} f(x) \left(F\psi \mathbf{1}_{\xi_d = \varphi(\xi)}\right)^\vee (x)\, dx$$

where (using pullback)

$$\begin{aligned}\left(F\psi \mathbf{1}_{\xi_d = \varphi(\xi)}\right)^\vee (x) &= \int_{\mathbb{R}^d} F(\tilde{\xi}) \psi(\tilde{\xi}) \mathbf{1}_{\xi_d = \varphi(\xi)} \mathrm{e}^{2\pi i x \cdot \tilde{\xi}}\, d\tilde{\xi} = \int_{\mathbb{R}^{d-1}} F(\xi) \psi(\xi) \mathrm{e}^{2\pi i (x' \cdot \xi + x_d \varphi(\xi))}\, d\xi \\ &= \int_S F(\xi) \mathrm{e}^{2\pi i x \cdot \xi}\, d\sigma(\xi) = (F d\sigma)^\vee(x)\end{aligned}$$

with the inconsistent notation $x' = (x_1, ..., x_{d-1}) \in \mathbb{R}^{d-1}$. This clarifies the computation in (2.2).

Remark 2.1. Had we started with a set of the form

$$S_\lambda := \{\xi \in \mathbb{R}^d : a(\xi) = \lambda\}$$

for a function $a : \mathbb{R}^d \to \mathbb{R}$ with

$$\nabla a(\xi) \neq 0 \quad \text{for } \xi \in a^{-1}(\Lambda)\,, \Lambda \subseteq \mathbb{R}\,,$$

then we define the measure on S_λ by the equality

$$d\Sigma_\lambda(\xi) = \frac{d\sigma_\lambda(\xi)}{|\nabla a(\xi)|}$$

where $d\sigma_\lambda(\xi)$ is the euclidean (Lebesgue) surface measure on S_λ. We remark that $d\Sigma_\lambda$ is sometimes also called the *canonical measure associated to a* (which is not intrinsic to S_λ, however), see also Strichartz [172, p. 705]. In particular, the elementary volume $d\xi$ in $\mathbb{R}^d$ satisfies

$$d\xi = d\lambda d\Sigma_\lambda(\xi)\,.$$

Moreover, by the implicit function theorem, the equation $a(\xi)=\lambda$ for λ close to some $\lambda_0\in\Lambda\subseteq\mathbb{R}$ defines a function $\xi_d=F(\xi',\lambda)$ for ξ close to $\xi^{(0)}\in S_{\lambda_0}$. Since the euclidean surface measure is given by $d\sigma_\lambda(\xi)=(1+|\nabla_{\xi'}F(\xi',\lambda)|^2)^{1/2}\,d\xi'$ (as we have seen above), we have

$$d\Sigma_\lambda(\xi)=\frac{d\xi'}{|\partial a(\xi)/\partial\xi_d|}\,.$$

Let us see the advantage of the introduction of $d\Sigma_\lambda(\xi)$. If one defines the Fourier multiplier $H_0=\mathcal{F}^*A\mathcal{F}$, where A is multiplication by the symbol $a(\xi)$ and $X\subseteq\mathbb{R}$ is some Borel set, then it is well known that its spectral projection is given by

$$E_0(X)=\mathcal{F}^*\mathbf{1}_{\{a^{-1}(X)\}}\mathcal{F}\,.$$

Thus, be the above discussion, we have

$$\langle\psi,E_0(X)\psi\rangle=\int_{a^{-1}(X)}|\hat\psi(\xi)|^2\,d\xi=\int_X d\lambda\int_{S_\lambda}|\hat f(\xi)|^2\,d\Sigma_\lambda(\xi)\,.$$

In particular, for a given measurable function $F:[0,\infty)\to\mathbb{R}$, we have

$$\langle\psi,F(H_0)\psi\rangle=\int_{\mathbb{R}_+}d\lambda\;F(\lambda)\int_{S_\lambda}|\hat\psi(\xi)|^2\,d\Sigma_\lambda(\xi)\,.$$

3. Necessary conditions

In this section, we will derive two common necessary conditions on p and q such that the extension estimate $\|(Fd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)}\lesssim_{p,q,d}\|F\|_{L^{q'}(S)}$, i.e., $R_S^*(q'\to p')$, holds. The *restriction conjecture* asserts that these two conditions are in fact also *sufficient*. The conjecture has been solved for the paraboloid and the sphere in two dimensions, and for the cone in up to four dimensions, but see also [185, Figures 1 and 2] for a more detailed `[and probably out of date]` summary of progress on this problem. In fact, the restriction problems for the three surfaces are related. Let us merely mention that the restriction conjecture of the sphere would imply the conjecture for the paraboloid since one can parabolically rescale the sphere to approach the paraboloid, but see also Tao [181] (where the surprising fact that the Bochner–Riesz conjecture implies the restriction conjecture is shown).

3.1. **The trivial condition.** By setting $F\equiv1$, we immediately see that we need $(d\sigma)^\vee\in L^{p'}(\mathbb{R}^d)$. In the case of the sphere and the paraboloid (which have non-vanishing Gaussian curvature), stationary phase computations yield

$$|(d\sigma)^\vee(x)|\lesssim(1+|x|)^{-(d-1)/2}\,,$$

i.e., we need $p'>2d/(d-1)$, respectively $p<2d/(d+1)$. For the sphere, an explicit computation using the Fourier–Bessel transform yields $(d\sigma)^\vee=2\pi|\xi|^{(2-d)/2}J_{(d-2)/2}(2\pi|\xi|)$. On the other hand, the asymptotics for the cone are slightly different, giving the condition $p'>2(d-1)/(d-2)$.

3.2. **Knapp's example.** We will sketch this example [189, 172] only for the sphere and the paraboloid (more precisely its intersection with the d dimensional unit cube). Assume that $R\gg1$ and take any interior point ξ_0 of the surface S. By a Taylor expansion, one sees that S contains a "cap" $\kappa\subseteq S$ centered at ξ_0 whose diameter is roughly R^{-1}. The cap has surface measure $\sim R^{-(d-1)}$ and can be packed into a d dimensional disk D of diameter R^{-1} and thickness R^{-2} which is oriented perpendicular to the unit normal of S at ξ_0. Now, let $F=\mathbf{1}_\kappa$ be the characteristic function of the cap κ and T be the tube dual to D. This is the tube which is centered at the origin, aligned along the unit normal to S at ξ_0 with length $\sim R^2$ and thickness $\sim R$. By the uncertainty principle (see also Appendix D), $(Fd\sigma)^\vee$ has magnitude $\sim\sigma(\kappa)\sim R^{-(d-1)}$ on a large portion of T (since the phase function $\mathrm{e}^{ix\cdot\xi}$ is basically constant for $\xi \in D$ and $x \in T$) and decays rapidly outside of T. In particular, we have

$$\|(Fd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)} \gtrsim |T|^{1/p'} R^{-(d-1)} \sim R^{(d+1)/p'-(d-1)}\,.$$

On the other hand,

$$\|F\|_{L^{q'}(S,d\sigma)} \sim |\kappa|^{1/q'} \sim R^{-(d-1)/q'}\,.$$

Letting $R \to \infty$ thus leads to the second necessary condition

$$\frac{d+1}{p'} \leq \frac{d-1}{q}$$

for $R_S^*(q' \to p')$ to hold. (Note that the Fourier transform $\widehat{d\sigma}$ of the measure $d\sigma$ associated to $\mathbb{S}^{d-1}$ decays like $|x|^{-(d-1)/2}$, i.e., it is $L^{p'}$-bounded for any $p' > 2d/(d-1)$. Thus the conjecture says that this $L^{p'}$-boundedness also holds for $\widehat{Fd\sigma}$.)

One can formulate a Knapp counterexample for any smooth hypersurface. Of course, the obtained necessary conditions become stronger as the surface becomes flatter. In the extreme case where the surface is infinitely flat (e.g. when it is a hyperplane), there are no estimates. In fact, the function $g(x) := (1+|x_1|)^{-1}$ lies in L^p for any $p > 1$ but has an infinite Fourier transform on every point of the hyperplane $\{\xi \in \mathbb{R}^d : \xi_1 = 0\}$.

Hence, we have the following conjectures, which are in fact all equivalent to each other [137, Section 19.3].

Conjecture 3.1. $\|\widehat{g\,d\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^p(S)}$ *for* $q > 2d/(d-1)$ *and* $q = (d+1)p'/(d-1)$.

Conjecture 3.2. $\|\widehat{g\,d\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^\infty(S)}$ *for* $q > 2d/(d-1)$.

Conjecture 3.3. $\|\widehat{g\,d\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^q(S)}$ *for* $q > 2d/(d-1)$.

Proposition 3.4. *The above three conjectures are equivalent to each other.*

Proof. Clearly the third version implies the second by Hölder. Once one shows the converse (i.e., second implies third version), the equivalence between the first and second version follows from interpolation. Observe that if $q = 2d/(d-1)$ and $q = (d+1)p'/(d-1)$, then $p = q$. For the rest, see [137, Theorem 19.8]. □

Before we come to the last example, we elaborate a bit on the situation of the paraboloid and perform some explicit computations for the reader's convenience. In fact, we will have a first encounter with "wave packets", an important tool that we will discuss in further detail in Section 9.

Knapp's example for the paraboloid - an explicit computation. Let F be a smooth, non-negative function with supp $F \subseteq \{|\xi| < 0.1\}$ and $\|F\|_1 = 1$. For $|\underline{x}|, |x_d| < 1$ the integral defining $\mathrm{Re}\,((Fd\sigma)^\vee(x))$ has no cancellation (since the phase function is strictly positive in this case), and hence $|(Fd\sigma)^\vee(x)|$ is nearly as large as possible, i.e.,

$$|(Fd\sigma)^\vee(x)| \geq \int_{\mathbb{R}^d} \cos(x\cdot(\underline{\xi}, |\underline{\xi}|^2))F(\xi)\,d\xi \geq \int_{\mathbb{R}^d} \cos(0.1+0.01)F(\xi)\,d\xi \sim 1\,.$$

For large $|x|$, the integrand oscillates rapidly in ξ, leading to cancellation in the integral, and hence a small contribution, i.e., $|(Fd\sigma)(x)| \ll 1$ for $|x| \gg 1$.

We will now rescale the above F such that "it lives on the paraboloid" by defining

$$F^R_{\xi_0,x_0}(\xi) = R^{d-1}\mathrm{e}^{2\pi i x_0(\underline{\xi},|\underline{\xi}|^2)}\varphi(R(\xi-\xi_0))$$

for some $R \gg 1$ where R^{-1} denotes the frequency scale of the parabolic subset (before it was the disk D)

$$\kappa_{\xi_0}^R := \{\xi \in \mathbb{P} : 0 \leq (\xi - \xi_0) \cdot \nu_{\xi_0} < 0.01R^{-2}\}$$

(centered at ξ_0) of the paraboloid $\mathbb{P}$, as before. Here, $\nu_{\xi_0} = (-2\xi_0, 1)$ denotes the upward normal to $\mathbb{P}$ at ξ_0. It is pretty clear that $\kappa_{\xi_0}^R$ is contained in a $R^{-1} \times \cdots \times R^{-1} \times R^{-2}$ rectangle centered at ξ_0 and whose short side is oriented along ν_{ξ_0}. We finally note that, due to the additional phase factor, $(F_{\xi_0,x_0}^R d\sigma)^\vee$ is going to be concentrated around x_0 in real space.

By scaling, the extension of this almost characteristic function on the inflated cap $\kappa_{\xi_0}^R$ is given by

$$(F_{\xi_0,x_0}^R d\sigma)^\vee(x) = \mathrm{e}^{2\pi i(x-x_0)\xi_0}(Fd\sigma)(R^{-1}((\underline{x} - \underline{x}_0) + 2(x-x_0)_d\xi_0), R^{-2}(x-x_0)_d)\,.$$

By the estimates on $(Fd\sigma)^\vee$, we see that $(F_{\xi_0,x_0} d\sigma)^\vee \sim 1$ on the tube

$$T_{\xi_0,x_0}^R = \{x \in \mathbb{R}^d : |(\underline{x} - \underline{x}_0) + 2(x-x_0)_d\xi_0| < R\,, |x_d - (x_0)_d| < R^2\}$$

which is centered at x_0, has width R and length R^2, and is aligned along ν_{ξ_0}. Off this tube, $(F_{\xi_0,x_0}^R d\sigma)^\vee$ decays rapidly. This shows that $\|(F_{\xi_0,x_0}^R d\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)} \gtrsim |T|^{1/p'} \sim R^{(d+1)/p'}$ whereas $\|F_{\xi_0,x_0}^R\|_{L^{q'}(\mathbb{P})} \sim R^{(d-1)/q}$ which again shows that $(d+1)/p' \leq (d-1)/q$ is necessary.

For $T = T_{\xi_0,x_0}^R$ and $F_T = F_{\xi_0,x_0}^R$, the extension $(F_{\xi_0,x_0}^R d\sigma)^\vee$ is called a *wave packet* associated to T. For any $R \gg 1$, a partition of unity directly decomposes the original function F as a sum of (unmodulated) R-caps, indexed by a collection of $\mathcal{O}(R^{d-1})$ tubes, i.e.,

$$F = \sum_T c_T F_T\,, \quad T = T_{\xi_0,0}^R\,.$$

Most of the coefficients c_T are of order $R^{-(d-1)}$ by scaling, the rest of them are even smaller. The curvature of $\mathbb{P}$ implies that distinct tubes T, T' with directions ν_T and $\nu_{T'}$ are separated by at least R^{-1} since

$$\text{angle} \sim \sin(\text{angle}) = \frac{R^{-2}}{R^{-1}} = R^{-1}\,.$$

3.3. **Hardy–Littlewood majorant conjecture.** The derivation of this conjecture is similar to the first (trivial) condition. Assume that F is a smooth function on S such that $\|F\|_{L^\infty(S)} \leq 1$. Since $Fd\sigma$ is pointwise dominated by $d\sigma$, it seems intuitive that $(Fd\sigma)^\vee$ should be "smaller" than $(d\sigma)^\vee$, too. The conjecture then states that the necessary conditions of the trivial condition should in fact be sufficient to obtain $R_S^*(\infty \to p')$ for completely general sets. It is known that the conjecture is true when p' is an even integer (using Plancherel's theorem) but is false for other values of p'. However, it may still be that the majorant conjecture is true if the set S is "non-pathological", e.g. in the cases for the sphere or the paraboloid.

3.4. **Are there restriction estimates for the plane?** We already mentioned in the introduction, that curvature was crucial for Stein's discovery of restriction estimates. Conversely, we may ask what happens when the curvature is zero, i.e., are there restriction estimates for the plane? Let us consider the hyperplane $\{\xi_d = 0\}$, or even only the subset

$$S = \{\xi \in \mathbb{R}^d : \xi_d = 0, |\xi| \leq 1\}$$

with the obvious surface measure $d\sigma$. Thus, by the Hölder and the Hausdorff–Young inequality, we have $\|\hat{f}\|_{L^q(S)} \lesssim \|f\|_p$ for $p = 1$ and arbitrary $q \geq 1$. However, these are the only estimates available.

Proposition 3.5. *Suppose $\|\hat{f}\|_{L^q(S)} \lesssim \|f\|_p$ holds for all test functions f and the above S. Then one must have $p = 1$.*

Proof. The idea is to consider functions whose Fourier transform is concentrated on and near S. For this, let $\psi \in \mathcal{S}(\mathbb{R}^d)$ with $\hat{\psi} \sim 1$ near the origin and let

$$f(x_1, ..., x_d) = \psi(x_1, ..., x_{d-1}, x_d/\lambda)$$

for large λ. Then $\|f\|_p \sim \lambda^{1/p}$ and

$$\hat{f}(\xi_1, ..., \xi_d) = \lambda \hat{\psi}(\xi_1, ..., \xi_{d-1}, \lambda\xi_d)\,.$$

In particular, $\hat{f} \sim \lambda$ on the pancake with dimensions $1 \times ... \times 1 \times \lambda^{-1}$ and $\hat{f}|_S \sim \lambda$. Thus, $\|\hat{f}\|_{L^q(S)} \sim \lambda$ and for $\|\hat{f}\|_{L^q(S)} \lesssim \|f\|_p$, i.e., $\lambda \lesssim \lambda^{1/p}$ to hold, we must have $p = 1$. □

In summary, there are no non-trivial restriction estimates for planes, even if we only consider compact pieces. The reason for this failure is that the plane is so flat that one can easily find $\hat{f}$ which are extremely large on and close to the plane.

3.5. **Curved surfaces, Fourier transforms of measures.** We just saw that there are no non-trivial restriction estimates for pieces of flat planes. Obviously, one may ask "how much restriction is possible, if we bend the plane a bit"?

As a starting point, suppose S has dimension $d-1$ and has non-vanishing Gauss curvature at every point. By that we mean the following. Let ξ_0 be any point of S and consider a rotation and translation of S such that ξ_0 becomes the origin and that the tangent plane to S at ξ_0 becomes the hyperplane $\xi_d = 0$. Then, near the origin at least, S can be given as a graph

$$\xi_d = \varphi(\xi_1, ..., \xi_{d-1})$$

where $\varphi \in C_c^\infty(\mathbb{R}^{d-1})$ and $\varphi(0) = \nabla\varphi(0) = 0$. Now, consider the $(d-1) \times (d-1)$ Hessian of φ, i.e.,

$$\left(\frac{\partial^2 \varphi}{\partial \xi_j \partial \xi_k}\right)_{j,k}(0)\,.$$

Its eigenvalues $\nu_1, ..., \nu_{d-1}$ are called the *principle curvatures of S at ξ_0*. The determinant of the Hessian, i.e., $\prod_{j=1}^{d-1} \nu_j$ is called the *Gaussian curvature of S at ξ_0*. Then, the following decay estimate for the Fourier transform of the surface measure $d\sigma$ of S can be proven via a stationary phase argument, see Stein [167, Chapter XIII, Section §3 and §5.7].

Proposition 3.6. *Suppose S is a smooth hypersurface in $\mathbb{R}^d$ with associated surface measure $d\sigma$. Assume $\psi \in C_c^\infty(\mathbb{R}^d)$ is a fixed function whose support intersects S in a compact subset of S and let $d\mu = \psi d\sigma$. If S has nowhere vanishing Gaussian curvature, then*

$$|(d\mu)^\vee(x)| \lesssim |x|^{-(d-1)/2}\,. \tag{3.1}$$

Remarks 3.7. (1) Herz [105] showed that, for $\psi = 1$, the surface need only be $C^{[(d-1)/2+2]}$ to obtain the above estimate on $|(d\mu)^\vee(x)|$. If the surface is $C^{[(d-1)/2+4]}$, he obtained the leading coefficient of the asymptotic expansion for $|(d\mu)^\vee(x)|$ as $|x| \to \infty$. If ψ is not constant on S, one can show (following the arguments of [167, Chapter XIII, §3.1] that $\psi \in C_c^{\lceil d/2 \rceil+2}$ and that the surface is $C^{\lceil d/2 \rceil+4}$ are sufficient conditions to obtain the above estimate on $|(d\mu)^\vee(x)|$.

(2) For smooth hypersurfaces where only k of the $d-1$ principal curvatures are non-vanishing, Littman [133] showed $|(d\mu)^\vee(x)| \lesssim |x|^{-k/2}$. (This is to be compared with the assertion in Stein [167, Chapter XIII, §3.2], where it is shown that $|(d\mu)^\vee(x)| \lesssim |x|^{-1/k}$ for hypersurfaces vanishing to k-th order, i.e., $\varphi(\xi_1, ..., \xi_{d-1}) = \mathcal{O}(|\xi|^k)$. (One says that S has finite type $k \in \{2, 3, ...\}$.)

Proposition 3.8 ([167, Chapter XIII, §3.2, Theorem 2])**.** *Suppose S is a smooth m-dimensional ($1 \leq m \leq d-1$) manifold in $\mathbb{R}^d$ of finite type. Let $d\mu = \psi d\sigma$ be as above. Then*

$$|(d\mu)^\vee(x)| \lesssim |x|^{-1/k}$$

where k is the type of S inside the support of ψ.

Using a T^*T argument (with T being the restriction operator) and the Hardy–Littlewood–Sobolev inequality, it is possible to establish the following (far from sharp) restriction estimate for finite type hypersurfaces.

Theorem 3.9. *Suppose S is a smooth m-dimensional ($1 \leq m \leq d-1$) manifold in $\mathbb{R}^d$ of type k. Then, one has $R_S(p \to 2)$ for any $1 \leq p \leq p_0$ with $p_0 = 2dk/(2dk-1)$.*

Although the theorem is not sharp, its main idea, namely exploiting cancellations through L^2 estimates, is the basis of the proof of the Tomas–Stein theorem.

Proof. If $f \in L^p(\mathbb{R}^d)$, then the $L^p(\mathbb{R}^d) \to L^2(S, d\sigma)$-boundedness of the restriction operator $\mathcal{R}_S$ is equivalent to the $L^p \to L^{p'}$-boundedness of $\mathcal{E}_S\mathcal{R}_S$ where $\mathcal{E}_S$ is the extension operator which is dual to $\mathcal{R}_S$. In particular, it suffices to show (cf. (4.10))

$$|\langle f, \mathcal{E}_S\mathcal{R}_S f\rangle| \lesssim \|f\|_p^2 .$$

Using the definition of $\mathcal{R}_s$ and $\mathcal{E}_S$, namely $(\mathcal{R}_S f)(\xi) = \hat{f}(\xi)|_S$ and $(\mathcal{E}_S g)(x) = \int_S \mathrm{e}^{ix\cdot\xi} g(\xi) d\sigma(\xi)$, we have $\langle f, \mathcal{E}_S\mathcal{R}_S f\rangle = \int \overline{f(x)} K(x-y) f(y)\, dx\, dy$ (which also equals $\langle \hat{f}, \hat{f} d\sigma\rangle = \langle f, f * (d\sigma)^\vee\rangle \leq \|f\|_p \|f * (d\sigma)^\vee\|_{p'} \leq \|f\|_p^2 \|(d\sigma)^\vee\|_{L^{p'/2,\infty}}$ using the weak Young inequality) where

$$K(x-y) = \int_S \mathrm{e}^{i\xi\cdot(x-y)} d\sigma(\xi) = (d\sigma)^\vee(x-y) .$$

Since $|(d\sigma)^\vee(x)| \lesssim |x|^{-1/k}$, we have $|\langle f, \mathcal{E}_S\mathcal{R}_S f\rangle| \lesssim \|f\|_p^2$ by the Hardy–Littlewood–Sobolev lemma if $p = 2dk/(2dk-1)$, i.e., $p' = 2dk$. The assertion follows from interpolation with $p = 1$. □

Extending the idea of Subsection 3.4, one can establish the following necessary condition for surfaces vanishing to k-th order. This argument generalizes the Knapp example, see Subsection 3.2. The original "Knapp" condition is obviously restored for $k = 2$.

Proposition 3.10. *Suppose $\varphi(\xi_1, ..., \xi_{d-1}) = \mathcal{O}(|\xi|^k)$ for some $k \geq 2$. Then, $R_S(p \to q)$ is only possible if*

$$p' \geq \frac{d+k-1}{d-1} q .$$

Proof. Let ψ be as in Proposition 3.5, i.e., $\psi \in \mathcal{S}(\mathbb{R}^d)$ with $\hat{\psi} \sim 1$ near the origin and let

$$f(x_1, ..., x_d) = \psi(x_1/\lambda^{1/k}, ..., x_{d-1}\lambda^{1/k}, x_d/\lambda)$$

for some large λ, i.e., f is a bump function on the $\lambda^{1/k} \times \cdots \lambda^{1/k} \times \lambda$ tube. In particular, $\|f\|_{L^p} \sim \lambda^{((d-1)/k+1)/p}$ and $\hat{f}(\xi) = \lambda^{(d-1)/k+1} \psi(\lambda^{1/k}\xi_1, ..., \lambda^{1/k}\xi_{d-1}, \lambda\xi_d)$. Since the volume of the cap where $\hat{f}$ does not decay rapidly is roughly $\lambda^{-(d-1)/k}$, we have $\|\hat{f}|_S\|_{L^q(S)} \gtrsim \lambda^{(d-1)/k+1} \times \lambda^{-(d-1)/(kq)}$. Comparing this with $\|f\|_p \sim \lambda^{(d-1+k)/p}$, yields the claimed necessary condition. □

The last two results and in particular Subsection 3.4, i.e., the absence of non-trivial restriction estimates on planes (which corresponds to the limit $k \to \infty$) underline the importance of curvature in restriction theory.

4. The Tomas–Stein restriction theorem

As far as positive results (besides the trivial $L^1 \to L^\infty$ estimate) go, we only have the following theorem of Tomas [189, 190] and Stein (1975, unpublished and [164]) which says that the restriction conjecture is indeed true for $q = 2$.

In fact, Stein gave two proofs of the restriction theorem. The first one relies on Tomas' (two pages long!) observation and on an extension of the classic Riesz–Thorin interpolation which is unpublished. We will discuss this in more detail in the second subsubsection. The other one establishes a theory of non-homogeneous oscillatory integral operators [164] that we will discuss in the next subsubsection. We emphasize that this approach uses ideas of Carleson and Sjölin [44] who proved the restriction theorem for $d = 3$ and $1 \leq p \leq 4/3$. The reader who is interested in the history prior to the Tomas–Stein theorem is invited to consult Tomas' paper [189].

Theorem 4.1. *If $1 \leq p \leq 2(d+1)/(d+3)$, then $R_S(p \to 2)$ holds.*

Remark 4.2. Bak and Seeger [3] extended the Tomas–Stein estimate to treat measures μ that satisfy

$$\sup_{\mathrm{rad}(B)\leq 1} \frac{\mu(B)}{\mathrm{rad}(B)^a} \leq A \tag{4.1}$$

and

$$\sup_{|\xi|\geq 1} |\xi|^b |\check{\mu}(\xi)| \leq C. \tag{4.2}$$

The number $\inf\{a : (4.1) \text{ holds for some } A < \infty\}$ is called the "dimension of μ", whereas the number $\inf\{a : (4.2) \text{ holds for some } C < \infty\}$ is called the "Fourier dimension". For (sufficiently) smooth hypersurfaces with non-vanishing Gauss curvature, one has $a = 2b = d - 1$. Bak and Seeger proved that the Tomas–Stein theorem extends to the stronger Lorentz-type estimate

$$\|\hat{f}\|_{L^2(S)} \lesssim_{d,a,b} A^{\frac{b}{d-a+b}} C^{\frac{d-a}{d-a+b}} \|f\|^2_{L^{p_c,2}(\mathbb{R}^d)} ,$$

where $p_c = 2(d-a+b)/(2(d-a)+b)$. One interesting application of this concerns surfaces where only k of the $d-1$ principal curvatures are non-vanishing. Littman [133] showed $|(d\mu)^\vee(x)| \lesssim |x|^{-k/2}$ for smooth hypersurfaces. In this case, $p_c = (2+k)/(2+k/2)$.

As we shall see, the proof heavily relies on the fact $q = 2$ and it has been very difficult (though not completely impossible) to push this argument beyond $q < 2$. We will now discuss the two approaches of the proof of this theorem.

4.1. Non-homogeneous oscillatory integral operators.

Following [161, Chapter 2] and Stein [167, Sections IX.1 and IX.2], the first approach consists in establishing a robust theory of non-homogeneous oscillatory integral operators of the form

$$I(\lambda) := \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda\varphi(y)} a(y)\, dy\, .$$

If φ has a non-degenerate critical point (i.e., $\nabla\varphi(y_0) = 0$ but $\det(\partial^2\varphi/\partial y_i \partial y_j) \neq 0$ when $y = y_0$), say at $y_0 = 0$, and a is a smooth cutoff function having small support, one can easily check that

$$|I(\lambda)| \sim \lambda^{-d/2} \quad \text{as } \lambda \to +\infty\, ,$$

whenever $a(0) \neq 0$, see, e.g., [161, Theorem 1.1.4]. The situation can be naturally extended by considering operators of the form

$$(T_\lambda f)(x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda\varphi(x,y)} a(x,y) f(y)\, dy\, , \quad \lambda > 0$$

where a is now a smooth cutoff function and $\varphi \in C^\infty(\mathbb{R}^m \times \mathbb{R}^d)$ is real. One may then, e.g., ask whether $T_\lambda f$ belongs to some L^p. The most basic result occurs when $m = d$. If φ is non-degenerate in the sense that the *mixed Hessian* satisfies the non-degeneracy condition

$$\det\left(\frac{\partial^2\varphi}{\partial x_j \partial x_k}\right) \neq 0\,,$$

then we shall find that

$$\|T_\lambda f\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-d/2}\|f\|_{L^2(\mathbb{R}^d)}\,.$$

This result obviously has the same flavor as the estimates for $I(\lambda)$, and, in fact, one can see that, for every λ, there are functions for which $\|T_\lambda f\|_2/\|f\|_2 \sim (1+\lambda)^{-d/2}$ if a is non-trivial.

However, there are many natural situations where the non-degeneracy condition is not met. The most popular example is of course $\varphi(x, y) = |x - y|$ for which the Hessian has only rank $d - 1$! The Tomas–Stein theorem will immediately follow from estimates on oscillatory integral operators with such phase functions.

4.1.1. *Non-degenerate oscillatory integral operators.* Let us however start with the simpler situation where the non-degeneracy condition is satisfied. The main theorem of this subsection is the following

Theorem 4.3. *Suppose φ is a real C^∞ phase function satisfying the non-degeneracy condition*

$$\det\left(\frac{\partial^2\varphi}{\partial x_j \partial x_k}\right) \neq 0 \tag{4.3}$$

on $\operatorname{supp}(a)$ *where $a \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d)$. Then for $\lambda > 0$,*

$$\left\|\int_{\mathbb{R}^d} \mathrm{e}^{i\lambda\varphi(x,y)} a(x,y) f(y)\, dy\right\|_{L^2(\mathbb{R}^d)} \lesssim \lambda^{-d/2}\|f\|_{L^2(\mathbb{R}^d)}\,. \tag{4.4}$$

If we let T_λ be the operator in (4.4), then clearly

$$\|T_\lambda f\|_\infty \lesssim \|f\|_1\,.$$

Thus, we obtain the following consequence by Riesz interpolation.

Corollary 4.4. *If $1 \le p \le 2$, then*

$$\left\|\int_{\mathbb{R}^d} \mathrm{e}^{i\lambda\varphi(x,y)} a(x,y) f(y)\, dy\right\|_{L^{p'}(\mathbb{R}^d)} \lesssim \lambda^{-d/p'}\|f\|_{L^p(\mathbb{R}^d)}\,. \tag{4.5}$$

Remark 4.5. Clearly, the phase function $\varphi(x, y) = \langle x, y\rangle$ leading to the standard Fourier transform satisfies the hypotheses of Theorem 4.3. Furthermore, (4.5) implies that

$$\left\|\int \mathrm{e}^{i\langle x,y\rangle} a(x/\sqrt{\lambda}, y/\sqrt{\lambda}) f(y)\, dy\right\|_{p'} \lesssim \|f\|_p\,,$$

i.e., (4.5) leads to another proof of the Hausdorff–Young inequality $\|\hat{f}\|_{p'} \lesssim \|f\|_p$.

Before we prove this theorem, we restate the non-degeneracy condition (4.3) in an equivalent form. Expanding

$$\nabla_x[\varphi(x,y) - \varphi(x,z)] = \left(\frac{\partial^2\varphi(x,y)}{\partial x_j \partial y_k}\right)(y - z) + \mathcal{O}(|y - z|^2)\,,$$

it is immediate that (4.3) is equivalent to

$$|\nabla_x[\varphi(x,y) - \varphi(x,z)]| \sim |y - z|\,, \quad |y - z| \ll 1\,. \tag{4.6}$$

This is the form that we shall use in the proof of Theorem 4.3.

Proof of Theorem 4.3. Using a smooth partition of unity, we can decompose $a(x,y)$ into a finite number of pieces each of which has the property that (4.6) holds on its support. Thus, we may assume without loss of generality that

$$|\nabla_x[\varphi(x,y)-\varphi(x,z)]| \gtrsim |y-z| \quad \text{on } \operatorname{supp}(a)$$

holds. The assertion then follows from Young's inequality for integral operators [161, Theorem 0.3.1], once we show that

$$|K_\lambda(y,z)| \lesssim_N (1+\lambda|y-z|)^{-N} \quad \text{for all } N\in\mathbb{N}$$

where

$$K_\lambda(y,z) = \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda[\varphi(x,y)-\varphi(x,z)]} a(x,y)\overline{a(x,z)}\,dx$$

is the integral kernel of $\langle f, T_\lambda f\rangle_{L^2(\mathbb{R}^d)}$. Since the above estimate just follows from a stationary phase argument (using (4.6)), we are already done. □

4.1.2. *Oscillatory integral operators related to the restriction theorem.* The main result in this subsection is that, under some natural additional geometric conditions on φ, we can prove that T_λ also maps $L^p(\mathbb{R}^{d-1})$ functions to $L^q(\mathbb{R}^d)$ functions with norm $\lambda^{-d/q}$ (see (4.5)).

As in the previous subsubsection, we will require a (modified) non-degeneracy condition of the form

$$\operatorname{rank}\left(\frac{\partial^2\varphi}{\partial y_j \partial z_k}\right) = d-1\,, \tag{4.7}$$

i.e., the mixed Hessian associated to the phase function has maximal rank. This condition alone would yield that $T_\lambda : L^p(\mathbb{R}^{d-1}) \to L^q(\mathbb{R}^d)$ is bounded with norm $\mathcal{O}(\lambda^{-(d-1)/q})$ if $q \geq 2$ and $p \geq q'$. To get the better result $\mathcal{O}(\lambda^{-d/q})$, we need an additional condition, more precisely, a *curvature hypothesis.*

To state it, we first notice that, since $\mathcal{C}_\varphi = \{(z, \varphi_z'(z,y), y, -\varphi_y'(z,y))\}$, (4.7), and the constant rank theorem imply that, for every $z_0 \in \operatorname{supp}_z(a)$, the image of $y \mapsto \varphi_z'(z_0,y)$, i.e.,

$$S_{z_0} = \prod_{T^*_{z_0}\mathbb{R}^d} (\mathcal{C}_\varphi) = \{\varphi_z'(z_0,y) : (z_0, \varphi_z'(z_0,y), y, -\varphi_y'(z_0,y)) \in \mathcal{C}_\varphi\}$$

is a C^∞ (immersed) hypersurface in $T^*_{z_0}\mathbb{R}^d$. Clearly, one can identify $T^*_{z_0}\mathbb{R}^d$ with $\mathbb{R}^d$. In this case, the curvature hypothesis says that

$$S_{z_0} \subseteq T^*_{z_0}\mathbb{R}^d \quad \text{has everywhere non-vanishing Gaussian curvature.} \tag{4.8}$$

Since changes of coordinates induce changes of coordinates in the cotangent bundle that are *linear* in the fibers, one concludes that (4.8) is (like (4.7)) an invariant condition. Notice that (4.7) is a condition involving second derivatives of the phase function whereas (4.8) is in fact a condition involving third derivatives.

If the two conditions (4.7) and (4.8) are met, we shall say that the phase function satisfies the *Carleson–Sjölin condition.* The main result of this subsubsection concerns estimates on oscillatory integral operators with such phase functions. It is due to Carleson and Sjölin [44] and Hörmander [108] in the two-dimensional case and to Stein [164, Theorem 10] in the higher-dimensional case.

Theorem 4.6. *Let T_λ as in* (4.4) *and suppose that the Carleson–Sjölin condition (i.e., the non-degeneracy condition* (4.7) *and the curvature condition* (4.8)*) holds. Then*

$$\|T_\lambda f\|_{L^q(\mathbb{R}^d)} \lesssim_p \lambda^{-d/q} \|f\|_{L^p(\mathbb{R}^{d-1})} \tag{4.9}$$

if $q = (d+1)p'/(d-1)$ and

(1) $1 \le p \le 2$ *for* $d \ge 3$*;*
(2) $1 \le p < 4$ *for* $d = 2$*.*

Bourgain [16] proved that the theorem can in fact not be improved beyond the range $1 \le p \le 2$ when $d \ge 3$. For the proof, we refer to [161, Theorem 2.2.1], see also Stein [164, Theorem 10] or [167, p. 380]. The details can also be found in Appendix A.2. Let us now actually see why the Tomas–Stein theorem is an immediate consequence of this theorem.

Corollary 4.7. *Suppose that* $S \subseteq \mathbb{R}^d$, $d \ge 2$ *is a* C^∞ *hypersurface with everywhere non-vanishing Gaussian curvature. Then, if* $d\sigma$ *is the Lebesgue measure on* S *and if* $d\mu = \beta d\sigma$ *with* $\beta \in C_c^\infty$*, it follows that*

$$\left(\int_S |\hat{f}(\xi)|^r \, d\mu(\xi)\right)^{1/r} \lesssim_S \|f\|_{L^s(\mathbb{R}^d)} \,,$$

provided that $r = (d-1)s'/(d+1)$ *and*
(1) $1 \le s \le 2(d+1)/(d+3)$ *for* $d \ge 3$*;*
(2) $1 \le s < 4/3$ *for* $d = 2$*.*

Notice that the exponents r and s are just conjugate to those in Theorem 4.6 which indicates that we will in fact prove the dual assertion, i.e., an extension estimate.

Proof. Without loss of generality, we may rotate and translate S such that $\xi_d = \varphi(\xi')$ for some $\varphi \in C^\infty(\mathbb{R}^{d-1})$ where as usual $\xi = (\xi', \xi_d) \in \mathbb{R}^d$. We shall now actually prove the (dual) extension estimate, i.e.,

$$\|(Fd\mu)^\vee\|_{L^{s'}(\mathbb{R}^d)} \lesssim \|F\|_{L^{r'}(S)}$$

where

$$\begin{aligned}(Fd\mu)^\vee(x) &= \int_S e^{2\pi i x\cdot\xi} F(\xi) d\mu(\xi) \\ &= \int_{\mathbb{R}^{d-1}} e^{2\pi i(x'\cdot\xi' + x_d\varphi(\xi'))} F(\xi', \varphi(\xi'))\beta(\xi', \varphi(\xi'))(1 + |\nabla\varphi(\xi')|^2)^{1/2} \, d\xi'\end{aligned}$$

and we used the pullback formula

$$d\mu(\xi) = \beta(\xi', \varphi(\xi'))(1 + |\nabla\varphi(\xi')|^2)^{1/2} d\xi' \equiv \Psi(\xi', \varphi(\xi')) d\xi'$$

To apply the Carleson–Sjölin theorem, we merely need to verify the non-degeneracy condition of Theorem 4.6. But this is easy since the Hessian of φ has rank $d-1$ and the curvature hypothesis holds by assumption. Thus, Theorem 4.6 implies $\|T_\lambda\|_{L^p(\mathbb{R}^{d-1}) \to L^q(\mathbb{R}^d)} \lesssim \lambda^{-d/q}$ where T_λ is defined by

$$(T_\lambda F)(x) := \int_{\mathbb{R}^{d-1}} e^{i\lambda\langle x, (\xi', \varphi(\xi'))\rangle} a(x, \xi') F(\xi', \varphi(\xi')) \cdot \Psi(\xi', \varphi(\xi')) \, d\xi' \,.$$

By scaling $x \mapsto x/\lambda$, this means that if p and q are as in Theorem 4.6, we have

$$\left\| \int_{\mathbb{R}^{d-1}} e^{2\pi i\langle x, (\xi', \varphi(\xi'))\rangle} a(x/\lambda, \xi') F(\xi', \varphi(\xi'))\Psi(\xi', \varphi(\xi')) \, dy \right\|_{L_x^q(\mathbb{R}^d)} \lesssim_p \|F(\cdot, \varphi(\cdot))\Psi(\cdot, \varphi(\cdot))\|_{L^p(\mathbb{R}^{d-1})}$$

for every $\lambda > 0$. Using once more the pullback formula, we conclude

$$\left\| \int_S e^{2\pi i x\cdot\xi} F(\xi) \, d\mu(\xi) \right\|_{L_x^q(\mathbb{R}^d)} \lesssim_p \|F\|_{L^p(S)} \,,$$

thereby showing the assertion. ☐

4.2. **The original arguments of Tomas and Stein.** Following Tao [179, Lecture 2], we will now outline the genesis of the Tomas–Stein theorem. In particular, we will encounter three basic interpolation theorems which are vital tools in (harmonic) analysis in general.

Squaring the desired restriction estimate shows that we need to prove

$$\int_S |\hat{f}(\xi)|^2 \, d\sigma(\xi) \lesssim \|f\|_p^2 \, .$$

We rewrite the left side as the L^2 inner product, use the convolution theorem and Hölder's inequality to obtain

$$\int_S |\hat{f}(\xi)|^2 \, d\sigma(\xi) = \langle \hat{f}, \hat{f} d\sigma \rangle = \langle \hat{f}, \mathcal{F}[f * \check{d\sigma}] \rangle = \langle f, f * \check{d\sigma} \rangle \leq \|f\|_p \|f * \check{d\sigma}\|_{p'} \, .$$

Thus, it suffices to prove

$$\|f * \check{d\sigma}\|_{p'} \lesssim \|f\|_p \, . \tag{4.10}$$

Note that this is just the TT^* method in disguise (i.e., showing that an operator T is $L^p \to L^2$-bounded is equivalent to showing that T^*T is $L^p \to L^{p'}$-bounded). The above observation was first made by Fefferman and Stein [72, p. 33ff].

We will now outline three proofs of (4.10).

4.2.1. *First attempt: fractional integration.* The most obvious tool to attack (4.10) would of course be to use the Hardy–Littlewood–Sobolev inequality (which is a special case of the "weak Young inequality").

Lemma 4.8. *If* $0 < \alpha < d$, $1 < p, q < \infty$, *and*

$$\frac{1}{q} + 1 = \frac{1}{p} + \frac{\alpha}{d} \, ,$$

then

$$\|f * |\cdot|^{-\alpha}\|_q \lesssim \|f\|_p$$

The other tool that we shall use is an interpolation theorem for weak-type operators. It turns out that the assumption that an operator has weak-type can be relaxed even a bit more. Recall that, for some measure spaces X, Y, a linear operator $T : X \to Y$ is said to have *weak-type* (p, q) if

$$|\{x \in X : |(Tf)(x)| > \lambda\}| \lesssim \lambda^{-q} \|f\|_p^q \quad \text{for all } f \in L^p, \lambda > 0 \, .$$

One can weaken this by considering only characteristic functions. We say that T has *restricted weak-type* (p, q) if

$$|\{x \in X : |(T\mathbf{1}_E)(x)| > \lambda\}| \lesssim \lambda^{-q} |E|^{p/q} \quad \text{for all } E \subset X, \lambda > 0 \, . \tag{4.11}$$

It is convenient to rephrase this estimate in a more symmetric form.

Lemma 4.9. *Suppose* $1 < p, q < \infty$. *Then* T *has restricted weak-type* (p, q) *if and only if*

$$|\langle T\mathbf{1}_E, \mathbf{1}_F \rangle| \lesssim |E|^{1/p} |F|^{1/q'} \tag{4.12}$$

for all sets $E \subseteq X$, $F \subseteq Y$.

As a comparison, recall that, by duality, the strong-type (p, q) estimate is equivalent to

$$|\langle Tf, g \rangle| \lesssim \|f\|_p \|g\|_{q'}$$

for all $f \in L^p$, $g \in L^{q'}$.

Proof. For our purposes, the necessity of the restricted weak-type estimate suffices which is why we will only deal with this direction. (For the other direction, one applies (4.12) to the set $F = \{\mathrm{Re}(T\mathbf{1}_E) > \lambda\}$.)

Using the layer cake representation and Fubini, we have

$$|\langle T\mathbf{1}_E, \mathbf{1}_F\rangle| \leq \int_F |T\mathbf{1}_E(x)|\,dx = \int_F \int_0^\infty \mathbf{1}_{\{|T\mathbf{1}_E|>\lambda\}}(x)\,d\lambda\,dx = \int_0^\infty |\{x \in F : |T\mathbf{1}_E(x)| > \lambda\}|\,d\lambda\,.$$

Recalling the restricted weak-type hypothesis, the integrand can be estimated by

$$|\{x \in F : |T\mathbf{1}_E(x)| > \lambda\}| \leq \min\{|F|, \lambda^{-q}|E|^{q/p}\}\,.$$

Thus,

$$|\langle T\mathbf{1}_E, \mathbf{1}_F\rangle| \lesssim \int_0^\infty \min\{|F|, \lambda^{-q}|E|^{q/p}\}\,d\lambda \lesssim |E|^{1/p}|F|^{1/q'}$$

by an elementary computation. □

Let us recall now

Lemma 4.10 (Marcinkiewicz interpolation)**.** *Suppose $1 < p_0 < q_0 < \infty$, $1 < p_1 < q_1 < \infty$, $p_0 < p_1$, $q_0 < q_1$, and T is of restricted weak-types (p_0, q_0) and (p_1, q_1). Then, T is of strong-type (p_θ, q_θ) for any $\theta \in (0,1)$ where $1/p_\theta = (1-\theta)/p_0 + \theta/p_1$ and similarly for q_θ.*

Proof. See Tao's notes [179, Lecture 2, Lemma 2.3] or Grafakos [96, Theorem 1.4.19] and Tao [186, Lecture 1, Lemma 8.5] for a further enhanced version. □

Using the decay estimate (3.1) (from stationary phase)

$$|\check{d\sigma}(x)| \lesssim |x|^{-(d-1)/2}\,,$$

and the Hardy–Littlewood–Sobolev inequality (Lemma 4.8) with $1/p' + 1 = 1/p + 2/p'$ and $2/p' = (d-1)/(2d)$, i.e., $p' = 4d/(d-1)$ and $p = 4d/(3d+1)$, we get

$$\|f * \check{d\sigma}\|_{p'} \lesssim \|f * |\cdot|^{-(d-1)/2}\|_{p'} \lesssim \|f\|_p\,.$$

In other words, we just proved the restriction estimate $R_S(4d/(3d+1) \to 2)$. By interpolation with the trivial estimate $R_S(1 \to 2)$, we thus get $R_S(p \to 2)$ for any $1 \leq p \leq 4d/(3d+1)$. This is a non-trivial statement, however, it is far from the best possible. Recall that the restriction conjecture says that p can go up to the endpoint $2(d+1)/(d+3)$.

The reason why we did not get a good estimate here is because we only performed pure size estimates, i.e., we merely exploited the decay of the convolution kernel $\check{d\sigma}(x)$. However, due to taking a Fourier transform, $\check{d\sigma}(x)$ actually also oscillates, in particular for *large* x. For instance, we have

$$\check{d\sigma}(x) = \text{const}\ J_{(d-2)/2}(|x|)/\sqrt{|x|}$$

for the sphere by the Fourier–Bessel transform. (Recall that $|J_\nu(x)| \lesssim |x|^{-1/2}$ for $|x| \to \infty$.) In $d = 3$, this reduces to

$$\frac{\sin(|x|)}{|x|}\,.$$

Crudely estimating these formulae by $|x|^{-1}$ is very inefficient.

4.2.2. *Second attempt: real interpolation.* In [189] Tomas introduced a very simple argument that made use both of the decay *and* the oscillations of the kernel $\check{d\sigma}$. This allowed him to get within an ε of the sharp result. The idea was to *decompose* $\check{d\sigma}$ *dyadically.* This idea is a very effective technique in harmonic analysis — break up your functions or kernels into many pieces in such a way that the behavior close to the singularity or at infinity can be treated very precisely, i.e., choose very small dyadic pieces where you need to obtain precise estimates. This approach works quite well, except when one has to recombine, i.e., to glue, all the pieces back together. In this way one often loses an ε, but rarely does one lose more than this.

Let us start with the main idea, i.e., the dyadic decomposition of $\check{d\sigma}$. For this, let φ be a radial bump function which equals 1 near 0 and is compactly supported. Then, define

$$\psi_k(x) := \varphi(2^{-k}x) - \varphi(2^{-k+1}x) .$$

Thus, ψ_k has size roughly 1 and is supported on the annulus $|x| \sim 2^k$. Moreover, the ψ_k are all related to each other by

$$\psi_k(x) = \psi_0(2^{-k}x)$$

and we have the telescopic identity

$$1 = \varphi(x) + \sum_{k>0} \psi_k(x) .$$

Thus, one can break up $f * \check{d\sigma}$ as

$$f * \check{d\sigma} = f * (\varphi\check{d\sigma}) + \sum_{k>0} f * (\psi_k\check{d\sigma}) .$$

Now, we may just use the triangle inequality, obtain

$$\|f * \check{d\sigma}\|_{p'} \leq \|f * (\varphi\check{d\sigma})\|_{p'} + \sum_{k>0} \|f * (\psi_k\check{d\sigma})\|_{p'} ,$$

and estimate each term separately. Note that one can (and should usually) be more sophisticated than the triangle inequality and use almost orthogonality results such as the Cotlar–Stein lemma (for operator norm bounds) or Carbery's lemma [37] (for Schatten norm bounds).

Since $d\sigma$ is a compactly supported measure, $\check{d\sigma}$ is a C^∞ function (it's complex analytic in fact). Thus, $\varphi\check{d\sigma} \in C_c^\infty$, i.e., the first term can be bounded by a constant times $\|f\|_p$ by Young's inequality.

Next, our goal is to estimate

$$\|f * (\psi_k\check{d\sigma})\|_{p'} \lesssim 2^{-\varepsilon k}\|f\|_p \tag{4.13}$$

since this would sum up nicely. We will prove such an estimate by interpolating between and $L^1 \to L^\infty$ and an $L^2 \to L^2$ estimate. The latter is just the one that will capture the oscillations of $d\sigma$!

Obtaining the $L^1 \to L^\infty$ estimate is easy because of the decay of the kernel $\check{d\sigma}$ and the fact that ψ_k localizes to the region $|x| \sim 2^k$. We obtain

$$\|f * (\psi_k\check{d\sigma})\|_\infty \lesssim \|f\|_1\|\psi_k\check{d\sigma}\|_\infty \lesssim \|f\|_1 2^{-k(d-1)/2} . \tag{4.14}$$

For the $L^2 \to L^2$ estimate, we use Plancherel and obtain

$$\|f * (\psi_k\check{d\sigma})\|_2 = \|\hat{f} \cdot (\psi_k\check{d\sigma})\|_2 \lesssim \|f\|_2\|\hat{\psi_k} * d\sigma\|_\infty .$$

Since ψ_k is smooth and compactly supported and acts as a mollifier, we have the standard estimate

$$|\hat{\psi_k}(\xi)| \lesssim \frac{2^{dk}}{< 2^k\xi >^N}$$

for any $N \in \mathbb{N}$. Thus, we obtain by scaling

$$|\hat{\psi}_k * d\sigma(\xi)| = \left| \int_S \hat{\psi}_k(\xi - \eta) d\sigma(\eta) \right| \lesssim 2^k$$

since we are integrating over S, i.e., an $d-1$-dimensional subset of $\mathbb{R}^d$ and therefore

$$\|f * (\psi_k \check{d\sigma})\|_2 \lesssim 2^k \|f\|_2 \,. \tag{4.15}$$

Interpolating (using Riesz–Thorin) between (4.14) and (4.15) thus yields

$$\|f * (\psi_k \check{d\sigma})\|_{p'} \lesssim 2^{-\varepsilon} \|f\|_p$$

for some ε, provided $p < 2(d+1)/(d+3)$.

Thus, by exploiting oscillation (via the Fourier transform-based $L^2 \to L^2$ estimate) *and* decay, we get $R_S(p \to 2)$ for all $1 \le p < 2(d+1)/(d+3)$. This is almost, but not quite, the sharp result as we are still missing the endpoint.

4.2.3. *Last attempt: complex interpolation.* In 1975 Stein (unpublished) obtained the endpoint estimate $R_S(2(d+1)/(d+3) \to 2)$ by extending the classic Riesz–Thorin interpolation "by adding a single letter to the alphabet" [75, p. 3]. Besides that, we will refuse to give in to the triangle inequality as we did in the last section and we will also make a special assumption on the localizing function ψ.

Theorem 4.11 (Stein's interpolation theorem)**.** *Assume T_z is an operator depending analytically on z in the strip $0 \le \mathrm{Re}\ z \le 1$. Suppose T_z is $L^{p_0} \to L^{q_0}$-bounded for* Re $z = 0$ *and $L^{p_1} \to L^{q_1}$-bounded for* Re $z = 1$. *Then T_θ is $L^{p_\theta} \to L^{q_\theta}$-bounded for $1/p_\theta = (1-\theta)/p_0 + \theta/p_1$, $1/q_\theta = (1-\theta)/q_0 + \theta/q_1$, and $\theta \in [0,1]$.*

Let $p = 2(d+1)/(d+3)$ and recall that we want to prove

$$\| \sum_{k>0} f * (\psi_k \check{d\sigma})\|_{p'} \lesssim \|f\|_p \,.$$

At this endpoint, (4.13), i.e.,

$$\|f * (\psi_k \check{d\sigma})\|_{p'} \lesssim 2^{-\varepsilon k} \|f\|_p \,,$$

only holds when $\varepsilon = 0$. In other words, to get the endpoint, we must not use the triangle inequality at this stage. We will therefore show the following two enhanced versions of the previous $L^1 \to L^\infty$ and $L^2 \to L^2$ bounds, namely

$$\| \sum_{k>0} 2^{[\frac{d-1}{2} + it]k} f * (\psi_k \check{d\sigma})\|_\infty \lesssim \|f\|_1 \tag{4.16a}$$

$$\| \sum_{k>0} 2^{[-1+it]k} f * (\psi_k \check{d\sigma})\|_2 \lesssim \|f\|_2 \tag{4.16b}$$

for all $t \in \mathbb{R}$. These two estimates, together with Theorem 4.11, then yield the desired estimate. (Note that, if we were only dealing with a fixed k, the above two estimates just correspond to (4.14) and (4.15)!) Let us now prove (4.16a) and (4.16b) and begin with the former. Rewriting it as

$$\|f * \sum_{k>0} 2^{[\frac{d-1}{2} + it]k} (\psi_k \check{d\sigma})\|_\infty \lesssim \|f\|_1 \,,$$

we see that it suffices (by Young's inequality) to prove

$$\| \sum_{k>0} 2^{[\frac{d-1}{2} + it]k} (\psi_k \check{d\sigma})\|_\infty \lesssim 1 \,.$$

But this just follows from the decay estimate $|\check{d\sigma}(x)| \lesssim |x|^{-(d-1)/2}$ since ψ_k localizes onto the dyadic region $|x| \in [2^k, 2^{k+1}]$, i.e.,

$$\sum_{k>0} 2^{[\frac{d-1}{2}+it]k}\psi_k(x) = \mathcal{O}(|x|^{(d-1)/2}) .$$

Note how we are being more efficient here than in the proof of (4.14).

Now let us turn to (4.16b). By the same arguments as in the previous section (i.e., Plancherel and Hölder), it suffices to prove

$$\| \sum_{k>0} 2^{[-1+it]k}(\hat{\psi}_k * d\sigma)\|_\infty \lesssim 1 .$$

Ignoring the cancellation coming from the 2^{itk} factor (which would be helpful however for $k \gg 1$), we will obtain this from

$$\sum_{k>0} 2^{-k}|(\hat{\psi}_k * d\sigma)(x)| \lesssim 1 . \tag{4.17}$$

In the previous section we already estimated $|(\hat{\psi}_k * d\sigma)(x)| \lesssim 2^k$, which is however just not good enough for our purpose. Instead, we shall establish the more sophisticated estimate

$$|(\hat{\psi}_k * d\sigma)(x)| \lesssim \begin{cases} 2^k(2^k d(x,S))^{-N} & \text{for } d(x,S) \geq 2^{-k} \\ 2^k + 2^k(2^k d(x,S)) & \text{for } d(x,S) \leq 2^{-k} \end{cases}$$

where $d(x,S) = |1-|x||$ is the distance of x to the unit sphere. Once we have this estimate, (4.17) follows from a routine calculation. Our task is thus to estimate

$$\left| \int_{\mathbb{S}^{d-1}} \hat{\psi}_k(x-\omega)d\sigma(\omega)\right| .$$

For $d(x,S) \geq 2^{-k}$ the claimed estimate follows, e.g., from the rapid decay $|\hat{\psi}_k(x)| \lesssim 2^{kd}(1 + 2^k|x|)^{-\tilde{N}}$ (possibly with $\tilde{N} \geq N + d$), decomposing $\mathbb{S}^{d-1}$ into regions where $d(x,\omega) \sim 2^{k+j}$ for some $j \geq 0$, and then summing in j.

Now, let us look at the region $d(x,S) \leq 2^{-k}$. If we just use the size estimate $|\hat{\psi}_k(x)| \lesssim 2^{kd}(1 + 2^k|x|)^{-\tilde{N}}$, we will end up with a bound of order $\mathcal{O}(2^k)$ which is just not good enough. Instead, we shall impose and exploit some moment conditions on ψ_k.

We first observe the Lipschitz bound

$$|\nabla(\widehat{\psi_k} * d\sigma)(x)| = 2^k|((2^{-k}\nabla\widehat{\psi_k}) * d\sigma)(x)| \lesssim 2^{2k}$$

where we used that also $2^{-k}\nabla\widehat{\psi_k}$ satisfies the above size estimate (by scaling) and that $d\sigma$ is supported on a $d-1$-dimensional manifold. Thus, for $y \in \mathbb{S}^{d-1}$,

$$(\widehat{\psi_k} * d\sigma)(x) - (\widehat{\psi_k} * d\sigma)(y) + (\widehat{\psi_k} * d\sigma)(y) \leq (\widehat{\psi_k} * d\sigma)(y) + d(x,y)(\nabla\widehat{\psi_k} * d\sigma)(z) \lesssim 2^k + 2^{2k}d(x,y) .$$

Thus, it suffices to consider points x on the unit sphere. By rotational symmetry, we may assume $x = e_n$, i.e., we need to show

$$| \int_{\mathbb{S}^{d-1}} \widehat{\psi_k}(e_n - \omega)d\sigma(\omega)| = \mathcal{O}(1) .$$

Because of the rapid decay of $\widehat{\psi_k}$ we may as well restrict ourselves to the region, say, $|e_n - \omega| < 1/10$. In this case, we parameterize $\omega \in \mathbb{S}^{d-1}$ as

$$\omega = (\underline{\omega}, \sqrt{1-|\underline{\omega}|^2}) , \quad \underline{\omega} \in \mathbb{R}^{d-1} .$$

Since we restricted our attention to $|e_n - \omega| < 1/10$, this means, it suffices to consider $|\underline{\omega}| < \sqrt{1-(9/10)^2} \ll 1$. Thus, we will estimate in the following

$$\int_{|\underline{\omega}|\ll 1} \widehat{\psi_k}(\underline{\omega}, 1-\sqrt{1-\underline{\omega}^2}) J(\underline{\omega})\, d\underline{\omega}$$

where $J(\underline{\omega})$ is the Jacobian appearing from our parameterization of ω. We may now rewrite this as a constant times

$$\int_{\mathbb{R}^{d-1}} \widehat{\psi_k}(\underline{\omega}, \mathcal{O}(\underline{\omega}^2))(1+\mathcal{O}(\underline{\omega}^2))\, d\underline{\omega} \tag{4.18}$$

modulo extremely tiny errors. We claim that this is quantity is

$$\int_{\mathbb{R}^{d-1}} \widehat{\psi_k}(\underline{\omega}, 0)\, d\underline{\omega} + \mathcal{O}(1)\,. \tag{4.19}$$

If this were the case, then we can simply choose φ, and thus ψ_0, so that

$$\int_{\mathbb{R}^{d-1}} \widehat{\psi_0}(\underline{\omega}, 0)\, d\underline{\omega} = 0$$

and this will achieve the desired estimate.

To prove the claimed approximation, we first observe that

$$\widehat{\psi_k}(\underline{\omega}, \mathcal{O}(\underline{\omega}^2)) = \widehat{\psi_k}(\underline{\omega}, 0) + \mathcal{O}\left(\frac{2^{(d+1)k}\underline{\omega}^2}{(1+2^k|\underline{\omega}|)^N}\right)$$

for all $N > 0$ by the rapid decay of $\widehat{\psi_k}$ and the mean value theorem. Thus, the error between (4.18) and (4.19) is at most

$$\int_{\mathbb{R}^{d-1}} \mathcal{O}\left(\frac{2^{(d+1)k}\underline{\omega}^2}{(1+2^k|\underline{\omega}|)^N}\right) d\underline{\omega} = \mathcal{O}(1)$$

which follows by scaling.

4.3. **Complex interpolation once more.** We shall give one further proof of the Tomas–Stein theorem which, however, does not use the dyadic decomposition of the kernel $(d\sigma)^\vee$. The technique that we will outline here, is in particular useful to obtain "uniform" resolvent estimates such as $\|(Q(D)-z)^{-1}\|_{p\to p'} \lesssim 1$ uniformly in $z \in \mathbb{C}$ for $|z| \geq 1$, $\mathrm{Im}(z) \neq 0$, $Q(\xi) = -\xi_1^2 - \xi_2^2 - ... - \xi_j^2 + \xi_{j+1}^2 + ... + \xi_d^2$ and p such that $2/(d+1) \leq 1/p - 1/p' \leq 2/d$, see, e.g., Kenig–Ruiz–Sogge [120, Theorem 2.3] where one interpolates between the $L^2 \to L^2$ bounds of $\mathrm{e}^{\zeta^2}(\Gamma(d/2+\zeta))^{-1}(Q(D)-z)^\zeta$ for $\mathrm{Re}(\zeta) = 0$ and the $L^1 \to L^\infty$ bounds of $\mathrm{e}^{\zeta^2}(\Gamma(d/2+\zeta))^{-1}(Q(D)-z)^\zeta$ for $\mathrm{Re}(\zeta) \in [-(d+1)/2, -d/2]$.

Theorem 4.12. *Suppose S is a smooth hypersurface in $\mathbb{R}^d$ with non-zero Gaussian curvature. Then*

$$\left(\int_{S_0} |\hat{f}(\xi)|^2\, d\sigma(\xi)\right)^{1/2} \lesssim_{p,S_0} \|f\|_{L^p(\mathbb{R}^d)}$$

holds for each $f \in \mathcal{S}(\mathbb{R}^d)$, $1 \leq p \leq 2(d+1)/(d+3)$, whenever S_0 is an open subset of S with compact closure in S.

The proof can be found in Stein [164, Theorem 3]. A more detailed exposition can be found in Stein–Shakarchi [168, Chapter 8, Theorem 5.2].

Proof. Suppose $0 \leq \psi \in C_c^\infty(\mathbb{R}^d)$. It will then suffice to prove

$$\left(\int_S |\hat{f}(\xi)|^2 \psi(\xi)\, d\sigma(\xi)\right)^{1/2} \lesssim_{p_0} \|f\|_{L^p(\mathbb{R}^d)} \tag{4.20}$$

for $p_0 = 2(d+1)/(d+3)$, the other cases follow from interpolation [1]. By covering the support of ψ by sufficiently many small open sets, it will be enough to prove the restriction estimate when (after a suitable rotation and translation of coordinates) the surface S is represented (in the support of ψ) as the graph $\xi_d = \varphi(\xi')$. Now, with $d\mu = \psi d\sigma$, the usual Plancherel argument implies

$$\int_S |\hat{f}(\xi)|^2 d\mu(\xi) = \int_{\mathbb{R}^d} \overline{(Tf)(x)} f(x)\, dx$$

where $(Tf)(x) = (K * f)(x)$ with

$$K(x) = \int e^{2\pi i x\cdot\xi}\, d\mu(\xi)\,.$$

Thus, we are left to show the $L^{p_0} \to L^{p_0'}$ boundedness of the convolution kernel K. To do so, we consider the family of kernels

$$K_s(x) := \frac{e^{s^2}}{\Gamma(s/2)} \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi} |\xi_d - \varphi(\xi')|^{-1+s} \eta(\xi_d - \varphi(\xi'))\tilde{\psi}(\xi')\, d\xi \tag{4.21}$$

where $\eta \in C_c^\infty(\mathbb{R}^d)$ is a bump function sitting at the origin and we set $\tilde{\psi}(\xi') = \psi(\xi', \varphi(\xi'))(1 + |\nabla\varphi(\xi')|^2)^{1/2}$ so that

$$d\mu(\xi) = \psi(\xi) d\sigma(\xi) = (1 + |\nabla\varphi(\xi')|^2)^{1/2} \psi(\xi', \varphi(\xi')) d\xi'.$$

Now, the change of variables $\xi_d \mapsto \xi_d + \varphi(\xi')$ in the above integrals shows that it equals

$$\zeta_s(x_d) \int_{\mathbb{R}^{d-1}} e^{2\pi i (x'\cdot\xi' + x_d\varphi(\xi'))} \tilde{\psi}(\xi')\, d\xi' = \zeta_s(x_d) K(x)$$

with

$$\zeta_s(x_d) = \frac{e^{s^2}}{\Gamma(s/2)} \int_{\mathbb{R}} e^{2\pi i x_d\xi_d} |\xi_d|^{-1+s} \eta(\xi_d)\, d\xi_d\,.$$

(Note that we now only need to study a "classical" function $\zeta_s(x_d)$ and the "regularized" kernels K_s since $K_0(x) = K(x)$ and we shall interpolate between K_{1+it} and $K_{-d/2+it}$ for $t \in \mathbb{R}$.) So, first it is well known that ζ_s has an analytic continuation in s which is an entire function. Moreover, $\zeta_0 \equiv 1$ (by an integration by parts, setting $s = 0$, and applying the fundamental theorem of calculus using $\eta(0) = 1$) and $|\zeta_s(x_d)| \lesssim < x_d >^{-\mathrm{Re}(s)}$, where the real part of s remains bounded from below (see also Stein–Shakarchi [168, Chapter 8, Lemma 4.6]). From these facts it follows that K_s has an analytic continuation to an entire function of s (whose values are smooth functions of $x_1, ..., x_d$ of at most polynomial growth). Moreover, one concludes

(1) $K_0(x) = K(x)$,
(2) $K_{(1-d)/2+it}$ is $L^1 \to L^\infty$ bounded with $|K_{(1-d)/2+it}(x)| \leq |\zeta_{(1-d)/2+it}(x_d)||K_0(x)| \lesssim 1$ for all $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$, and
(3) K_{1+it} is L^2 bounded with $|\hat{K}_{1+it}(\xi)| \lesssim 1$ for all $\xi \in \mathbb{R}^d$ and $t \in \mathbb{R}$.

[1] In fact, the interpolation argument shows that we can take q so that the restriction estimate holds where the $L^2(S, d\sigma)$ norm is replaced by the $L^q(S, d\sigma)$ norm with $q = (d-1)p'/(d+1)$ which is the optimal relation between p and q.

In fact, (2) follows from the estimates $|K_0(x)| = |(d\mu)^\vee(x)| \lesssim |x|^{-(d-1)/2}$ and $|\zeta_{(1-d)/2+it}(x_d)| \lesssim |x_d|^{(d-1)/2}$ whereas (3) follows from the definition (4.21) of K_s. (The integrand in (4.21) is just $\hat{K}_s$ which is clearly bounded for $s = 1$.) Thus, we have shown that the analytic family T_s of operators, defined by $T_s f = K_s * f$ satisfies

$$\|T_{(1-d)/2+it}f\|_\infty \lesssim \|f\|_1\,, \quad t \in \mathbb{R}$$

because of (2) and also

$$\|T_{1+it}f\|_2 \lesssim \|f\|_2\,, \quad t \in \mathbb{R}$$

because of (3) and Plancherel. Thus, by complex interpolation ($0 = (1-\theta) + \theta\cdot(-d/2)$, i.e., $\theta = 2/(d+1)$, $1-\theta = (d-1)/(d+1)$, and $1/p = (1-\theta)/2 + \theta/1 = (d+3)/(2d+2)$, i.e., $p = 2(d+1)/(d+3) = p_0$), we obtain the asserted $L^{p_0} \to L^{p_0'}$ boundedness. □

4.4. **A final word on complex interpolation.** Frank and Sabin [87, Proposition 1] noticed that once one proves the $L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ boundedness of some operator T on $\mathbb{R}^d$ via complex interpolation, one not only obtains that W_1TW_2 is $L^2(\mathbb{R}^d)$ bounded for arbitrary $W_1, W_2 \in L^{2p/(2-p)}(\mathbb{R}^d)$ (by Hölder's inequality). In fact, W_1TW_2 also belongs to some trace ideal $\mathcal{S}^p(L^2(\mathbb{R}^d))$. We quote

Proposition 4.13 (Frank–Sabin [87, Proposition 1])**.** *Let T_z be an analytic family of operators in $\mathbb{R}^d$ in the sense of Stein defined on the strip $-\lambda_0 \leq \mathrm{Re} z \leq 0$ for some $\lambda_0 > 1$. Assume that the bounds*

$$\|T_{iy}\|_{2,2} \leq M_0 \mathrm{e}^{a|y|}\,, \quad \|T_{-\lambda_0+iy}\|_{1,\infty} \leq M_1 \mathrm{e}^{b|y|}\,, \quad \forall y \in \mathbb{R} \tag{4.22}$$

hold for some $a, b \geq 0$ and for some $M_0, M_1 \geq 0$. Then, for all $W_1, W_2 \in L^{2\lambda_0}(\mathbb{R}^d : \mathbb{C})$, it holds that $W_1T_{-1}W_2 \in \mathcal{S}^{2\lambda_0}(L^2(\mathbb{R}^d))$ with

$$\|W_1T_{-1}W_2\|_{\mathcal{S}^{2\lambda_0}(L^2(\mathbb{R}^d))} \leq M_0^{1-1/\lambda_0} M_1^{1/\lambda_0} \|W_1\|_{L^{2\lambda_0}(\mathbb{R}^d)} \|W_2\|_{L^{2\lambda_0}(\mathbb{R}^d)}\,. \tag{4.23}$$

The basic idea is to use complex interpolation between Schatten spaces (cf. Simon [156, Theorem 2.9]) applied to the holomorphic family $W_1^{-z}T_zW_2^{-z}$ for $\mathrm{Re}(z) \in [-\lambda_0, 0]$. One then interpolates between the $L^2 \to L^2$ bound and the Hilbert–Schmidt estimate

$$\|W_1^{\lambda_0-iy}T_{-\lambda_0+iy}W_2^{\lambda_0-iy}\|_{\mathcal{S}^2}^2 = \int_{\mathbb{R}^d} dx \int_{\mathbb{R}^d} dy\, |W_1(x)|^{2\lambda_0} |W_2(x)|^{2\lambda_0} |T_{-\lambda_0+iy}(x,y)|^2\,.$$

Sometimes, there are better estimates for $|T_{-\lambda_0+iy}(x,y)|$ available than a simple uniform bound. This may, e.g., be the case when T_z is a differential operator such as $(-\Delta-\zeta)^z$. Then one may use tools like the Hardy–Littlewood–Sobolev inequality and so on.

Proof. For $W_j = |W_j|\mathrm{e}^{i\varphi_j}$ we have

$$\|W_1T_{-1}W_2\|_{\mathcal{S}^{2\lambda_0}} \leq \|\mathrm{e}^{i\varphi_1}\|_{L^2\to L^2} \||W_1|\,T_{-1}\,|W_2|\|_{\mathcal{S}^{2\lambda_0}} \|\mathrm{e}^{i\varphi_2}\|_{L^2\to L^2} \leq \||W_1|\,T_{-1}\,|W_2|\|_{\mathcal{S}^{2\lambda_0}}\,.$$

Thus, we can reduce to the case where W_1, W_2 are non-negative. Moreover, by a density argument, we may suppose W_1, W_2 to be simple functions. For simple $W_1, W_2 \geq 0$ we now define the family of operators

$$S_z := W_1^{-z}T_zW_2^{-z}$$

which is still analytic in the sense of Stein in the strip $-\lambda_0 \leq \mathrm{Re}(z) \leq 0$ and satisfies $S_{-1} = W_1T_{-1}W_2$. On the right border of the strip, i.e., $\mathrm{Re}(z) = 0$, we have

$$\|S_{is}\|_{L^2\to L^2} \leq \|W_1^{-is}\|_\infty \|W_2^{-is}\|_\infty \|T_{is}\|_{L^2\to L^2} \leq M_0\mathrm{e}^{a|s|}\,, \quad s \in \mathbb{R}\,.$$

On the left border, we prove that $S_{-\lambda_0+is}$ is Hilbert–Schmidt. Indeed, we obtain

$$\|S_{-\lambda_0+is}\|_{\mathcal{S}^2}^2 = \int_{\mathbb{R}^d} dx\, dy\, W_1(x)^{2\lambda_0} W_2(y)^{2\lambda_0} |T_{-\lambda_0+is}(x,y)|^2 \leq M_1^2 \mathrm{e}^{2b|s|} \|W_1\|_{2\lambda_0}^{2\lambda_0} \|W_2\|_{2\lambda_0}^{2\lambda_0} .$$

Thus, by complex interpolation for Schatten ideals (cf. Simon [156, Theorem 2.9]), it follows that $S_{-1} \in \mathcal{S}^{2\lambda_0}(L^2(\mathbb{R}^d))$ with

$$\|S_{-1}\|_{\mathcal{S}^{2\lambda_0}(L^2(\mathbb{R}^d))} \leq M_0^{1-1/\lambda_0} M_1^{1/\lambda_0} \|W_1\|_{2\lambda_0} \|W_2\|_{2\lambda_0} .$$

This concludes the proof. □

If, in addition, the operator T_{-1} can be factorized in A^*A, we have the following duality principle which is interesting in the context of many-fermion systems, where a one-particle density matrix of orthonormal wave functions has the form

$$\gamma = \sum_j \nu_j |f_j\rangle\langle f_j|$$

for some $\nu_j \geq 0$ satisfying $\sum_j \nu_j = 1$ and orthonormal $f_j \in L^2(\mathbb{R}^d)$.

Lemma 4.14 (Frank–Sabin [87, Lemma 3])**.** *Let $\mathcal{H}$ be a separable Hilbert space. Assume that A is a $\mathcal{H} \to L^{p'}(\mathbb{R}^d)$ bounded operator for some $1 \leq p \leq 2$ and let $\alpha \geq 1$. Then the following are equivalent.*

(i) There is a constant $C > 0$ such that

$$\|WAA^*\overline{W}\|_{\mathcal{S}^\alpha(L^2(\mathbb{R}^d)))} \leq C\|W\|_{L^{2p/(2-p)}(\mathbb{R}^d)}^2 , \quad \forall W \in L^{2p/(2-p)}(\mathbb{R}^d : \mathbb{C}) . \tag{4.24}$$

(ii) There is a constant $C' > 0$ such that for any orthonormal system $(f_j)_{j\in J} \in \mathcal{H}$ and any sequence $(\nu_j)_{j\in J} \subseteq \mathbb{C}$,

$$\left\| \sum_{j\in J} \nu_j |Af_j|^2 \right\|_{L^{p'/2}(\mathbb{R}^d)} \leq C' \left(\sum_{j\in J} |\nu_j|^{\alpha'} \right)^{1/\alpha'} . \tag{4.25}$$

Moreover, the values of the optimal constants C and C' coincide.

Applications of these principles include

- Tomas–Stein restriction estimates in Schatten spaces [87, Theorems 2 (and 4) and 3 (and 5)] (the optimality of these results is shown in [87, Theorem 6]),
- Strichartz estimates for the paraboloid $S = \{(\omega, \xi) \in \mathbb{R} \times \mathbb{R}^d, \omega = -|\xi|^2\}$ (Schrödinger with $-\Delta$) [87, Theorems 7,8, and 9] , the cone $S = \{(\omega, \xi) \in \mathbb{R} \times \mathbb{R}^d, \omega^2 = |\xi|^2\}$ (wave, respectively Schrödinger with $\sqrt{-\Delta}$) [87, Theorem 10], and the two-sheeted hyperboloid $S = \{(\omega, \xi) \in \mathbb{R} \times \mathbb{R}^d, \omega^2 = 1 + |\xi|^2\}$ (Klein–Gordon, respectively Schrödinger with $\sqrt{1-\Delta}$) [87, Theorem 11], and
- uniform Sobolev inequalities à la Kenig–Ruiz–Sogge [120] for $-\Delta$ (see [87, Theorems 12 and 13] and Subsection 17.5 later) and Cuenin [53] for $(m^2 - \Delta)^{s/2} - m$ with $0 < s < d$ and $\sum_{j=1}^d \alpha_j(-i\nabla_j) + \beta m$ with $m \geq 0$.
- Eigenvalue estimates for Schrödinger operators with complex potentials [87, Theorem 16]. See also Frank–Laptev–Lieb–Seiringer [85], Frank [83, 84], Frank–Simon [88], Laptev–Safronov [127], Safronov [150].

For the sake of completeness we state the Tomas–Stein estimate for trace ideals here.

Theorem 4.15. *Let $d \geq 2$, $S \subseteq \mathbb{R}^d$ be a smooth, compact surface with non-zero Gaussian curvature. Then*

$$\|W_1 \mathcal{F}_S^* \mathcal{F}_S W_2\|_{\mathcal{S}^{\frac{(d-1)q}{d-q}}(L^2(\mathbb{R}^d))} \lesssim \|W_1\|_{L^{2q}(\mathbb{R}^d)} \|W_2\|_{L^{2q}(\mathbb{R}^d)}\,, \quad q \in [1, \frac{d+1}{2}]\,. \tag{4.26}$$

See also Theorem 18.4 for an alternative proof not relying on complex interpolation.

Sketch of the proof. By Proposition 4.13 and the proof of Theorem 4.12, one obtains the bound

$$\|W_1 \mathcal{F}_S^* \mathcal{F}_S W_2\|_{\mathcal{S}^{\frac{2p}{2-p}}(L^2(\mathbb{R}^d))} \lesssim \|W_1\|_{L^{\frac{2p}{2-p}}(\mathbb{R}^d)} \|W_2\|_{L^{\frac{2p}{2-p}}(\mathbb{R}^d)}\,, \quad p \in [1, \frac{2(d+1)}{d+3}]\,.$$

The assertion in the theorem follows by interpolating this (with $p = 2(d+1)/(d+3)$) with the trace bound

$$\|W_1 \mathcal{F}_S^* \mathcal{F}_S W_2\|_{\mathcal{S}^1(L^2(\mathbb{R}^d))} \leq \|W_1 \mathcal{F}_S^*\|_{\mathcal{S}^2(L^2(S), L^2(\mathbb{R}^d))} \|W_2 \mathcal{F}_S^*\|_{\mathcal{S}^2(L^2(S), L^2(\mathbb{R}^d))}$$

where

$$\|W \mathcal{F}_S^*\|_{\mathcal{S}^2(L^2(S, d\sigma), L^2(\mathbb{R}^d))}^2 = \int_{\mathbb{R}^d} |W(x)|^2 \int_S d\sigma(\xi) = \sigma(S) \|W\|_{L^2(\mathbb{R}^d)}^2\,.$$

Here we used that the integral kernel of $W\mathcal{F}_S^*$ is given by $W(x)\mathrm{e}^{2\pi i x \cdot \xi}$. ☐

The "trace class restriction theorem" 4.15 is predated by an observation of Birman, Koplienko, Krein, Kurda, and Yafaev (see also [11, 12] for asymptotic results on the eigenvalues of the scattering matrix, in particular of $F_S V F_S^*$ in $L^2(S)$ and [202, Proposition 8.1.3] for a textbook reference) in the context of scattering amplitudes. The proof uses the same strategy above by interpolating in Schatten ideals between a bounded operator (when the potential decays like $|x|^{-1-\varepsilon}$) and a trace class operator (when the potential decays like $|x|^{-d-\varepsilon}$).

Theorem 4.16 (Yafaev [202, Proposition 8.1.3])**.** *Suppose $|V(x)| \lesssim (1+|x|)^{-\rho}$ for some $\rho > 1$ and let*

$$\begin{aligned} \Gamma_0(\lambda) : \mathcal{S}(\mathbb{R}^d) &\to \mathcal{S}'(\mathbb{S}^{d-1}) \\ f &\mapsto 2^{-1/2} \lambda^{(d-2)/4} \hat{f}(\lambda^{1/2} \cdot) \end{aligned}$$

be the rescaled restriction operator on $\sqrt{\lambda}\mathbb{S}^{d-1}$ (see [202, Formula (1.2.5)]*) with adjoint Γ_0^* (extension operator) given by*

$$(\Gamma_0^* g)(x) = 2^{-1/2} \lambda^{(d-2)/4} \int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi i \sqrt{\lambda} x \cdot \xi} g(\xi)\, d\Sigma(\xi)$$

where $d\Sigma$ is the Lebesgue measure on $\mathbb{S}^{d-1}$ (see [202, Formula (1.2.7) or Proposition 8.1.3]*). Then for all $\lambda > 0$ and $p > (d-1)/(\rho-1)$ and $p \geq 1$, one has*

$$\|\Gamma_0 V \Gamma_0^*\|_{\mathcal{S}^p(L^2(\mathbb{S}^{d-1}))} \lesssim \lambda^{-1/2 + (d-1)/(2p)}.$$

4.5. **A simpler L^2-based restriction theorem.** Notice that the L^2 estimate in Tomas' arguments was based only on dimensionality considerations. This suggests that there should be an L^2 bound for $\widehat{fd\sigma}$ (similar to the classical trace lemma $\|\hat{g}\|_{L^2(S)} \lesssim \|g\|_{L^2_\sigma(\mathbb{R}^d)}$ for all $\sigma > 1/2$) valid under very general conditions.

Theorem 4.17. *Let ν be a positive finite measure satisfying the dimensional condition*[2]

$$\nu(B_x(r)) \leq C r^\alpha\,. \tag{4.27}$$

[2] We only require one half of the Ahlfors–David regularity condition which would involve also a matching lower bound.

Then there is a bound

$$\|\widehat{fd\nu}\|_{L^2(B_0(R))} \leq CR^{\frac{d-\alpha}{2}} \|f\|_{L^2(d\nu)} \,. \tag{4.28}$$

The proof relies on the following well known

Lemma 4.18 (Schur's test)**.** *Let (X,μ) and (Y,ν) be measure spaces and $K(x,y)$ a measurable function on $X \times Y$ satisfying*

$$\int_X |K(x,y)|\, d\mu(x) \leq A \quad \text{for all } y \in Y\,,$$
$$\int_Y |K(x,y)|\, d\nu(y) \leq B \quad \text{for all } x \in X\,.$$

Let $T_K : \mathcal{S}(X) \to \mathcal{S}'(Y)$ defined by $T_K f(x) = \int_Y K(x,y) f(y)\, d\nu(y)$. Then, for $f \in L^2(d\nu)$, the integral defining $T_K f$ converges $d\mu$-a.e. and we have

$$\|T_K f\|_{L^2(d\mu)} \leq \sqrt{AB} \|f\|_{L^2(d\nu)}\,.$$

Proof of Theorem 4.17. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ such that $\varphi(x) \geq 1$ on $B_0(1)$ and that $\hat{\varphi} \in C_c^\infty(\mathbb{R}^d)$. Denote the scaled version by $\varphi_R(x) = \varphi(x/R)$. Then,

$$\|\widehat{fd\nu}\|_{L^2(B_0(R))} \leq \|\varphi_R(x) \widehat{fd\nu}(-x)\|_{L^2(dx)} = \|\widehat{\varphi_R} * (fd\nu)\|_{L^2(\mathbb{R}^d)}\,.$$

We will now use Schur's test to estimate the operator norm of the convolution operator $\widehat{\varphi_R} * (\cdot)$. On the one hand, we have

$$\int_{\mathbb{R}^d} |R^d \hat{\varphi}(R(\xi - \eta))|\, d\xi = \|\hat{\varphi}\|_{L^1(\mathbb{R}^d)} < \infty$$

and on the other hand,

$$\int |R^d \hat{\varphi}(R(\xi-\eta))|\, d\nu(\eta) \lesssim R^{d-\alpha}$$

since $\hat{\varphi}$ was assumed to be compactly supported and the dimensional condition on $d\nu$. Thus, by Schur's test,

$$\|\widehat{fd\nu}\|_{L^2(B_0(R))} \lesssim R^{(d-\alpha)/2} \|f\|_{L^2(d\nu)}\,,$$

thereby establishing the claim. □

4.6. **Trace theorems.** We recall some classical trace theorems—originally due to Gagliardo [90]—from PDE or scattering theory and follow Yafaev [202, Section 1.1]. Throughout, we assume that $S \subseteq \mathbb{R}^d$ is a hypersurface, i.e., a codimension one manifold. We start with the case where S can be parameterized by, say, a continuous function $F : \Omega \to \mathbb{R}$ where $\Omega \subseteq \mathbb{R}^{d-1}$ is an open set, i.e., $\xi_d = F(\xi')$.

Proposition 4.19. *Let $\alpha > 1/2$, then*

$$\int_\Omega |\hat{u}(\xi', F(\xi')|^2\, d\xi' \lesssim \frac{1}{2\alpha - 1} \int_{\mathbb{R}^d} (1 + x_d^2)^\alpha |u(x)|^2\, dx \leq \frac{1}{2\alpha-1} \|u\|^2_{L^2_\alpha(\mathbb{R}^d)}\,, \tag{4.29}$$

where $\|u\|_{L^2_\alpha(\mathbb{R}^d)} = \|u\|_{L^2(\mathbb{R}^d,\, <x>^{2\alpha}\, dx)}$.

Proof. Let

$$\tilde{u}(\xi', x_d) := \int_{\mathbb{R}^{d-1}} \mathrm{e}^{-2\pi i x' \cdot \xi'} u(x', x_d)\, dx'$$

and

$$\hat{u}(\xi', F(\xi')) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i (x' \cdot \xi' + x_d F(\xi'))} u(x', x_d)\, dx = \int_{\mathbb{R}} \mathrm{e}^{-2\pi i x_d F(\xi')} \tilde{u}(\xi', x_d)\, dx_d\,.$$

Then, by Schwarz

$$|\hat{u}(\xi', F(\xi'))|^2 \lesssim \frac{1}{2\alpha-1} \int_{\mathbb{R}} < x_d >^{2\alpha} |\tilde{u}(\xi', x_d)|^2 \, dx_d$$

and therefore

$$\begin{aligned}\int_\Omega d\xi' \; |\hat{u}(\xi', F(\xi'))|^2 &\lesssim \frac{1}{2\alpha-1} \int_\Omega d\xi' \int_{\mathbb{R}} dx_d \; < x_d >^{2\alpha} |\tilde{u}(\xi', x_d)|^2 \\ &= \frac{1}{2\alpha-1} \int_{\mathbb{R}^d} < x_d >^{2\alpha} |u(x)|^2 \, dx\end{aligned}$$

where we used Plancherel in the prime variables in the final step. This concludes the proof. □

Remark 4.20. The proof also goes through when $\mathbb{R}^d$ in position space is replaced by $\mathbb{Z}^d$ and correspondingly $\mathbb{R}^d$ by $\mathbb{T}^d = [-1/2, 1/2]^d$ the d-dimensional torus in Fourier space. Of course one needs that S is actually contained in $\mathbb{T}^d$.

Next we recall that the left side of (4.29) is actually Hölder continuous with respect to a variation of the function F, i.e., with respect to perturbation of the surface in question.

Proposition 4.21. *Let $\alpha > 1/2$ and*

$$\theta = \begin{cases} \alpha - 1/2 & \text{for } \alpha < 3/2 \\ 1-\varepsilon & \text{for any } \varepsilon \in (0,1) \quad \text{for } \alpha = 3/2 \\ 1 & \text{for } \alpha > 3/2 \end{cases} .$$

Then

$$\int_\Omega |\hat{u}(\xi', F(\xi')) - \hat{u}(\xi', \tilde{F}(\xi'))|^2 \, d\xi' \lesssim_{\alpha,\theta} \sup_{\xi' \in \Omega} |F(\xi') - \tilde{F}(\xi')|^{2\theta} \int_{\mathbb{R}^d} (1 + x_d^2)^\alpha |u(x)|^2 \, dx \,. \quad (4.30)$$

Proof. We use the same representation as before, i.e.,

$$\tilde{u}(\xi', x_d) := \int_{\mathbb{R}^{d-1}} \mathrm{e}^{-2\pi i x' \cdot \xi'} u(x', x_d) \, dx'$$

which allows us to write for $\xi_d, \tilde{\xi}_d \in \mathbb{R}$,

$$\hat{u}(\xi', \xi_d) - \hat{u}(\xi', \tilde{\xi}_d) = \int_{\mathbb{R}} \left(\mathrm{e}^{-2\pi i x_d \xi_d} - \mathrm{e}^{-2\pi i x_d \tilde{\xi}_d} \right) \tilde{u}(\xi', x_d) \, dx_d$$

and estimate

$$\begin{aligned}|\hat{u}(\xi', \xi_d) - \hat{u}(\xi', \tilde{\xi}_d)|^2 &\le \left(\int_{\mathbb{R}} \sin^2(\pi(\tilde{\xi}_d - \xi_d) x_d) < x_d >^{-2\alpha} \; dx_d \right) \\ &\quad \times \left(\int_{\mathbb{R}} < x_d >^{2\alpha} |\tilde{u}(\xi', x_d)|^2 \, dx_d \right) .\end{aligned}$$

using Schwarz. The first integral on the right side is bounded by a constant times $|\xi_d - \tilde{\xi}_d|^{2\theta}$. Thus, setting $\xi_d = F(\xi')$ and $\tilde{\xi}_d = \tilde{F}(\xi')$ yields

$$|\hat{u}(\xi', \xi_d) - \hat{u}(\xi', \tilde{\xi}_d)|^2 \lesssim |F(\xi') - \tilde{F}(\xi')|^{2\theta} \int_{\mathbb{R}} < x_d >^{2\alpha} |\tilde{u}(\xi', x_d)|^2 \, dx_d \,.$$

Integrating this over ξ' and using Plancherel as in the proof of Proposition 4.19 concludes the proof. □

The integral in (4.29) is actually taken over the hypersurface S given by the equation $\xi_d = F(\xi')$. Clearly, the Lebesgue measure $d\xi'$ can be replaced by any scaled version $C(\xi')d\xi'$ for any $C(\xi') > 0$. If, e.g., $F \in C_b^1(\Omega)$ (i.e., F is continuously differentiable with bounded derivative), then one can integrate over the Euclidean measure

$$d\Sigma_S(\xi') = \left(1 + |\nabla F(\xi')|^2\right)^{1/2} d\xi' \tag{4.31}$$

on S. Thus, Proposition 4.19 implies that

$$\|\hat{u}\|_{L^2(S)}^2 \lesssim_\alpha \|u\|_{L^2_\alpha(\mathbb{R}^d)}^2 \,, \quad \alpha > 1/2 \,.$$

This inequality shows that sufficiently fast decaying functions have Fourier transforms that can be meaningfully restricted to hypersurfaces; more precisely, this shows

$$L^2_{1/2+\varepsilon}(\mathbb{R}^d) \hookrightarrow L^2(S) \,.$$

Moreover, for the existence of a trace of a function, it suffices to have decay only in some directions transversal to S. Moreover, the relation $L^2_{1/2+\varepsilon}(\mathbb{R}^d) \hookrightarrow L^2(S)$ can be generalized as follows.

Proposition 4.22. *Suppose a hypersurface can be covered by a finite number of hypersurfaces S_j given in their own coordinate systems by functions $\xi_d = F_j(\xi')$ where ξ' belong to open sets $\Omega_j \subseteq \mathbb{R}^{d-1}$. Assume further that $F_j \in C_b^1(\Omega_j)$ for all j. Then, one has*

$$\|\hat{u}\|_{L^2(S)}^2 \lesssim \frac{1}{2\alpha - 1} \|u\|_{L^2_\alpha(\mathbb{R}^d)}^2 \,, \quad \alpha > 1/2$$

and hence $L^2_{1/2+\varepsilon}(\mathbb{R}^d) \hookrightarrow L^2(S)$.

5. Randomized restriction in $\mathbb{Z}^2$

We follow Bourgain [20]. Let $\Gamma \subseteq \Pi^2$ be a smooth, compact hypersurface with nowhere vanishing Gaussian curvature. We could in principle work in any dimension if Γ satisfied these assumptions, however in the model case where Γ is the level set

$$\Gamma_\lambda = \{m(\xi) := 2(\cos(2\pi\xi_1) + \cos(2\pi\xi_2)) = \lambda\} \,, \quad |\lambda| \in (\tau, 4 - \tau) \,, \quad 0 < \tau \ll 1 \,,$$

this is only satisfied in $d = 2$. We denote by Σ_λ the arclength-measure of Γ_λ. Thus, by stationary phase,

$$|\widehat{d\Sigma_\lambda}(n)| \lesssim (1 + |n|)^{-1/2} \,, \quad n \in \mathbb{Z}^2$$

and so by Stein's proof (using complex interpolation), we have

Lemma 5.1. *Let μ be a measure supported on Γ_λ such that $\mu \ll \Sigma_\lambda$ and $d\mu/d\Sigma_\lambda \in L^2(\Gamma, d\Sigma)$. Then*

$$\|\hat{\mu}\|_{\ell^6(\mathbb{Z}^2)} \lesssim \left\|\frac{d\mu}{d\Sigma_\lambda}\right\|_{L^2(\Gamma, d\Sigma_\lambda)} .$$

Now suppose

$$V_\omega(n) := \omega_n |n|^{-\varepsilon} v(n)$$

where $\{\omega_n : n \in \mathbb{Z}^2\}$ are independent Bernoulli or normalized Gaussian random variables, and $v \in \ell^p(\mathbb{Z}^2)$ and some $p \geq 1$. By Hölder and the above Tomas–Stein estimate, we obtain the deterministic estimate

$$\|F_{S_{t_2}} v F_{S_{t_1}}^*\|_{L^2(\Gamma_{t_1}, d\Sigma_{t_1}), L^2(\Gamma_{t_2}, d\Sigma_{t_2})} \lesssim \|v\|_{\ell^{3/2}(\mathbb{Z}^2)} \,.$$

(As usual F_{S_t} and $F_{S_t}^*$ denote the Fourier restriction and extension operators with respect to $(\Gamma_t, d\Sigma_t)$.) However, the randomness of V_ω allows us to relax the decay condition on $v(n)$ substantially.

Theorem 5.2. *Let* $V_\omega(n) := \omega_n|n|^{-\varepsilon}v(n) = \sum_{\ell\geq 0} V_\ell$ *with* $v \in \ell^3(\mathbb{Z}^2)$*,* $V_0(n) = V\mathbf{1}_0(n)$ *and* $V_\ell(n) = V\mathbf{1}_{2^{\ell-1}\leq|n|\leq 2^\ell}(n)$*. Then*

$$\mathbb{E}_\omega\left[\|F_S V_\ell F_S^*\|_{L^2(\Gamma,d\Sigma),L^2(\Gamma,d\Sigma)}\right] \lesssim \|v\|_{\ell^3(\mathbb{Z}^2)}\cdot 2^{-c\ell}\,, \quad \text{some } 0<c<\varepsilon\,. \tag{5.1}$$

If $V_\omega(n) := \omega_n w(n)$ *with* $w \in \ell^{3-\varepsilon}(\mathbb{Z}^2)$*, then*

$$\mathbb{E}_\omega\left[\|F_{S_{t_2}} V_\ell F_{S_{t_1}}^*\|_{L^2(\Gamma_{t_1},d\Sigma_{t_1}),L^2(\Gamma_{t_2},d\Sigma_{t_2})}\right] \lesssim \|w\|_{\ell^{3-\varepsilon}(\mathbb{Z}^2)}\cdot 2^{-c\ell}\,, \quad \text{some } c=c(\varepsilon)>0\,. \tag{5.2}$$

To prepare the proof we collect some classic results in geometry of Banach spaces and probability theory. The crucial ingredients going into the proof of Theorem 5.2 are the "dual to Sudakov bound" (Theorem 5.7) and "Dudley's L^{ψ_2} estimate[3]" (Corollary 5.29). The proof will be concluded in Subsection 5.5.

The above statement continues to hold for potentials V_ω in $\mathbb{R}^d$.

Theorem 5.3. *Let* $d\in\mathbb{N}\setminus\{1\}$*,* $\{\omega_n : n\in\mathbb{Z}^d\}$ *be independent Bernoulli or normalized Gaussian random variables, and*

$$V_\omega(x) = \sum_{n\in\mathbb{Z}^d} v_n\omega(n)\mathbf{1}_{Q_n}(x)\,, \quad x\in\mathbb{R}^d$$

with $Q_n = [0,1)^d + n$ *and* $n\in\mathbb{Z}^d$*. Suppose that either* $(v_n)_{n\in\mathbb{Z}^d}\subseteq \ell^{d+1-\varepsilon}(\mathbb{Z}^d)$ *or* $v_n = |n|^{-\varepsilon}w_n$ *with* $(w_n)_{n\in\mathbb{Z}^d}\subseteq\ell^{d+1}(\mathbb{Z}^d)$*. Let* $\Gamma_t\subseteq\mathbb{R}^d$ *be a family of smooth and compact codimension one manifolds whose Gaussian curvatures never vanish and* F_{S_t} *and* $F_{S_t}^*$ *denote the corresponding Fourier restriction and extension operators. Then we have*

$$\mathbb{E}_\omega\|F_{S_{t_2}}V_\omega F_{S_{t_1}}^*\|_{L^2(\Gamma_{t_1},d\Sigma_{t_1}),L^2(\Gamma_{t_2},d\Sigma_{t_2})} \lesssim \min\{\|v\|_{d+1-\varepsilon},\|w\|_{d+1}\}\,.$$

The proof of this theorem follows the lines of that of Theorem 5.2 and will also be given in Subsection 5.5.

Since Tomas–Stein theorems also give rise to resolvent estimates (see, e.g., Cuenin [54]) there are various applications of this result that are discussed in Subsection 5.6.

5.1. **Facts in geometry of Banach spaces and entropy bounds.** We mainly follow Pajor and Tomczak–Jaegermann [144] and Bourgain–Lindenstrauss–Milman [24].

The main question we pursue here is the following: suppose we are given two subsets D and B of a linear space. What is the minimal number of dilated translates of B needed to cover D? I.e., for given $t>0$ we want to find good upper bounds on

$$N(D,B,t) := \min\{k\in\mathbb{N} : \exists (x_i)_{i=1}^k \text{ s.t. } D\subseteq \bigcup_{i=1}^k x_i + tB\}\,. \tag{5.3}$$

Sometimes we will use a slightly different terminology, e.g., in the following more concrete situation. Suppose that (T,d) is a compact metric space, then

$$N(T,d,\varepsilon) := \text{smallest number of } \varepsilon\text{-balls needed to cover } T\,. \tag{5.4}$$

Example 5.4. If T is the unit ball in an n-dimensional Banach space, such as $\ell_d^p \equiv (\mathbb{R}^d,\|\cdot\|_p)$, then

$$N(T,d,\varepsilon) \leq (1+2/\varepsilon)^n\,. \tag{5.5}$$

See, e.g., Figiel–Lindenstrauss–Milman [77, Lemma 2.4] or Bourgain–Lindenstrauss–Milman [24, Lemma 2.4].

[3] ψ_2 stands for the Orlicz function e^{y^2}.

The main estimate we are interested here is the "dual to Sudakov estimate" due to Pajor and Tomczak–Jaegermann [144]. We will recast their estimate in a different form that will be useful in our context and follow Bourgain et al [24].

Let $|||\cdot||| := [\cdot,\cdot]^{1/2}$ denote the euclidean norm and scalar product on $\mathbb{R}^n$ with B^n its unit ball and $\mathbb{S}^{n-1}$ its boundary, the euclidean sphere. Suppose $\|\cdot\|$ is another norm on $\mathbb{R}^n$ and denote by $X = (\mathbb{R}^n, \|\cdot\|)$ and $X^* = (\mathbb{R}^n, \|\cdot\|_*)$ the corresponding Banach space and its dual. Here,

$$\|\psi\|_* := \sup\{|[\psi,\varphi]| : \varphi \in X \text{ with } \|\varphi\|_X \leq 1\}.$$

Since all norms on $\mathbb{R}^n$ are equivalent to each other there exist $a, b > 0$ such that

$$a^{-1}|||x||| \leq \|x\| \leq b|||x|||. \tag{5.6}$$

If, e.g., $|||\cdot||| = \|\cdot\|_2$ and $\|\cdot\| = \|\cdot\|_1$, then $a = 1$ and $b = n^{1/2}$. By interpolation, we obtain $a = 1$ and $b = n^{1/p-1/2}$ if $\|\cdot\| = \|\cdot\|_p$ and $1 \leq p \leq 2$. By duality, we have $a = n^{1/2-1/p}$ and $b = 1$ for $p \in [2,\infty]$.

Next, define the median M_r of $r(x) := \|x\|$ on $\mathbb{S}^{n-1}$ by

$$\mu(\{x \in \mathbb{S}^{n-1} : r(x) \geq M_r\}) \geq \frac{1}{2} \quad \text{and} \quad \mu(\{x \in \mathbb{S}^{n-1} : r(x) \leq M_r\}) \geq \frac{1}{2} \tag{5.7}$$

where μ is the associated normalized, rotation invariant Haar measure on $\mathbb{S}^{n-1}$. Moreover, the average A_r of $r(x) = \|x\|$ on $\mathbb{S}^{n-1}$ is given by

$$A_r := \int_{\mathbb{S}^{n-1}} \|x\| d\mu(x). \tag{5.8}$$

We record the following lemma on the comparability of A_r and M_r.

Lemma 5.5. *If $b \leq \sqrt{n}$ in (5.6), then there is $C > 0$ such that $|A_r - M_r| < C$. If additionally $ab \leq \sqrt{n}$ in (5.6), then $1/2 \leq A_r/M_r \leq C$.*

Proof. See Milman–Schechtman [138, Lemma 5.1]. □

Next, we rewrite A_r using homogeneity and polar coordinates as

$$A_r = \frac{a_n}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \|x\| \exp\left(-\frac{|||x|||^2}{2}\right) dx, \quad a_n \sim n^{-1/2}. \tag{5.9}$$

A probabilistic way to write this is to consider n independent and normalized Gaussian variables $\{g_j(\omega)\}_{j=1}^n$ on some probability space (Ω, ρ). Then [24, (4.4)]

$$A_r = a_n \int_\Omega \left\| \sum_{j=1}^n g_j(\omega) e_j \right\| d\rho(\omega), \quad a_n \sim n^{-1/2} \tag{5.10}$$

where $\{e_j\}_{j=1}^n$ denotes an orthonormal basis in $\mathbb{R}^n$.

We are now in position to state two answers to the question posed at the beginning of this section. The first gives an upper bound on the minimal number of euclidean t-balls needed to cover B_X, the unit ball in $X = (\mathbb{R}^n, \|\cdot\|)$.

Proposition 5.6 (Sudakov [174]). *Let $X = (\mathbb{R}^n, \|\cdot\|)$ and $|||\cdot|||$ be the euclidean norm on $\mathbb{R}^n$. Let B_X and B^n denote the unit balls in $\mathbb{R}^n$ with respect to the norms $\|\cdot\|$, respectively $|||\cdot|||$. Then*

$$\log N(B_X, |||\cdot|||, t) = \log N(B_X, B^n, t) \leq C \cdot n \cdot \left(\frac{A_{r^*}}{t}\right)^2, \tag{5.11}$$

where $A_{r^*} := \int_{\mathbb{S}^{n-1}} \|x\|_* d\mu(x)$.

The following estimate is dual to that one.

Theorem 5.7 (Dual to Sudakov [144, 24])**.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ and $|||\cdot|||$ be the euclidean norm on $\mathbb{R}^n$. Let B_X and B^n denote the unit balls in $\mathbb{R}^n$ with respect to the norms $\|\cdot\|$, respectively $|||\cdot|||$. Then*

$$\log N(B^n, \|\cdot\|, t) = \log N(B^n, B_X, t) \leq C\cdot n \cdot \left(\frac{A_r}{t}\right)^2 . \tag{5.12}$$

Remark 5.8. It is useful to have another interpretation of $N(B^n, \|\cdot\|, t)$ in mind. It is precisely the minimal size of a finite subset $\mathcal{E} \subseteq B^n$ that satisfies

$$\max_{x\in B^n} \min_{x'\in\mathcal{E}} \|x - x'\| < t\,.$$

Proof. We follow [24, Proposition 4.2]. Let σ be the gaussian probability measure on $\mathbb{R}^n$ defined by

$$d\sigma(x) = \frac{1}{(2\pi)^{n/2}} \exp(-\frac{|||\cdot|||^2}{2})\,dx\,.$$

Then by

$$A_r = \frac{a_n}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \|x\| \exp(-\frac{|||\cdot|||^2}{2})\,dx = \int_{\mathbb{S}^{n-1}} \|x\| d\mu(x)$$

and Chebyshev's inequality

$$\sigma(\{x\in\mathbb{R}^n : |f(x)| > \alpha\}) \leq \alpha^{-p} \int\limits_{|f(x)|\geq\alpha} |f(x)|^p d\sigma(x)\,,$$

we have, for $f(x) = \|x\|$, $p = 1$, and $\alpha = 2A_r/a_n$ that

$$\sigma(\{x\in\mathbb{R}^n : \|x\| > \frac{2A_r}{a_n}\}) \leq \frac{a_n}{2A_r}\int_{\mathbb{R}^n} \|x\| d\sigma(x) = \frac{1}{2}$$

and so

$$\sigma(\{x\in\mathbb{R}^n : \|x\| \leq \frac{2A_r}{a_n}\}) \geq \frac{1}{2}\,. \tag{5.13}$$

Next, suppose $\{x_j\}_{j=1}^N$ is a maximal subset of B^n relative to the requirement that $\|x_j - x_\ell\| \geq t$ for all $j \neq \ell$. This ensures that the sets $\{x_j + \frac{t}{2}B_X\}_{j=1}^N$ have disjoint interior. Since σ is a probability measure on $\mathbb{R}^n$, this disjointness implies

$$1 \geq \sum_{j=1}^N \sigma(\{y_j + \frac{2A_r}{a_n}B_X\}) = N\sigma(\{y_j + \frac{2A_r}{a_n}B_X\})\,, \quad \text{where } y_j = \frac{4A_r}{a_n t}x_j\,. \tag{5.14}$$

By convexity of e^{-u^2} for $u > 0$ and symmetry of B_X with respect to the origin (for the first estimate in the following formula) and (5.13) (for the second estimate), we have for fixed $j = 1, ..., N$,

$$\begin{aligned}
\sigma(\{y_j + \frac{2A_r}{a_n}B_X\}) &= \frac{1}{(2\pi)^{n/2}} \int_{\frac{2A_r}{a_n}B_X} dx\, \exp\left(-\frac{|||x-y_j|||^2}{2}\right) \\
&\geq \frac{1}{(2\pi)^{n/2}} \int_{\frac{2A_r}{a_n}B_X} dx\, \exp\left(-\frac{|||x-y_j|||^2 + |||x+y_j|||^2}{4}\right) \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\frac{2A_r}{a_n}B_X} dx\, \exp\left(-\frac{|||x|||^2 + |||y_j|||^2}{2}\right) = \exp\left(-\frac{|||y_j|||^2}{2}\right)\cdot\sigma\left(\frac{2A_r}{a_n}B_X\right) \\
&\geq \frac{1}{2}\exp\left(-\frac{|||y_j|||^2}{2}\right) \geq \frac{1}{2}\exp\left(-\frac{4A_r^2}{(ta_n)^2}\right)\,.
\end{aligned}$$

Combining this with (5.14) then finally gives

$$N \leq 2\exp(\frac{4A_r^2}{(ta_n)^2}) \Rightarrow \log N \lesssim \frac{1}{n} \cdot \left(\frac{A_r}{t}\right)^2 .$$

This concludes the proof. □

In the following we are interested in finding large euclidean sections in a finite-dimensional normed space.

Definition 5.9. Let X, Y be two n-dimensional normed spaces. The *Banach–Mazur distance* betweeen them is defined as

$$d(X,Y) := \inf\{\|T\| \cdot \|T^{-1}\| : T : X \to Y \text{ isomorphism}\} . \tag{5.15}$$

If $d(X,Y) \leq \lambda$, we say that X and Y are λ-isomorphic.

Obviously, $d(X,Y) \geq 1$ and $d(X,Y) = 1$ if and only if X and Y are isometric. Thus, by the discussion after (5.6), we see that (cf. [138, p. 20])

$$d(\ell_n^2, \ell_n^p) \leq n^{|1/2-1/p|} \quad 1 < p < \infty .$$

Theorem 5.10 (F. John [138, Theorem 3.3])**.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be an n-dimensional normed space. Let D be the ellipsoid of maximal volume inscribed in B_X and $|||\cdot|||$ be the euclidean norm induced by D, i.e., $D = \{x \in B_X : |||x||| \leq 1\}$. Then*

$$n^{-1/2}|||x||| \leq \|x\| \leq |||x|||$$

and consequently $d(\ell_n^2, X) \leq \sqrt{n}$ (where ℓ_n^2 is equipped with $|||(x_1,...,x_n)|||^2 = \sum_{j=1}^n |x_j|^2$.)

5.2. **Connection between probability theory and geometry of Banach spaces.** We follow Milman–Schechtman [138].

Definition 5.11. Let X be a normed space and $\{\varepsilon_j\}_{j\in\mathbb{N}}$ be Rademacher signs. For $1 \leq p \leq 2 \leq q < \infty$, and $n \in \mathbb{N}$ we define the *type p* (resp. *cotype q*) constants $T_p(X,n)$ (resp. $C_q(X,n)$) of X as the smallest T (resp. C) such that

$$\left(\mathbb{E}\left\|\sum_{j=1}^n \varepsilon_j x_j\right\|^2\right)^{1/2} \leq T\left(\sum_{j=1}^n \|x_j\|^p\right)^{1/p}$$

resp.

$$\left(\sum_{j=1}^n \|x_j\|^q\right)^{1/q} \leq C\left(\mathbb{E}\left\|\sum_{j=1}^n \varepsilon_j x_j\right\|^2\right)^{1/2}$$

for all $x_1, ..., x_n \in X$. If $T_p(X) := \sup_n T_p(X,n) < \infty$ (resp. $C_q(X) := \sup_n C_q(X,n) < \infty$) we say that X *has type p resp. cotype q* with type p constant $T_p(X)$ and cotype constant $C_q(X)$.

Theorem 5.12 (Kahane's inequality)**.** *Let X be a normed space and $p \in [1,\infty)$. Then there is a constant $K_p > 0$ such that*

$$\mathbb{E}\|\sum_{j=1}^n \varepsilon_j x_j\| \leq [\mathbb{E}(\|\sum_{j=1}^n \varepsilon_j x_j\|^p)]^{1/p} \leq K_p\,\mathbb{E}\|\sum_{j=1}^n \varepsilon_j x_j\| \tag{5.16}$$

where $x_1, ..., x_n \in X$ and $\{\varepsilon_j\}_{j\in\mathbb{N}}$ are Rademacher distributed.

Proof. See, e.g., [138, Theorem 9.2]. □

Example 5.13. L^p has type p and cotype 2 for $1 \leq p \leq 2$. Respectively, L^q has type 2 and cotype q for $2 \leq q < \infty$. (See [138, Example 9.3].) This follows from Kahane's inequality and Khinchine's inequality (for $0 < p < \infty$)

$$\|(\sum_j |x_j|^2)^{1/2}\|_{L^p}^p \sim \int \mathbb{E}(|\sum_j \varepsilon_j x_j|^p)\, dx\,. \tag{5.17}$$

Definition 5.14. A L^2-normalized, random Gaussian variable is a random variable $g(\omega)$ whose distribution is given by

$$\mathbb{P}(g(\omega) \leq t) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{t} \mathrm{e}^{-s^2/2}\, ds\,.$$

Let $\{g_j\}_{j=1}^\infty$ be a sequence of independent gaussian variables normalized in L^2. For a normed space X, $1 \leq p \leq 2 \leq q < \infty$, and $n \in \mathbb{N}$ we define the *gaussian type* p (resp. *cotype* q) constants $\alpha_p(X,n)$ (resp. $\beta_q(X,n)$) of X as the smallest T (resp. C) such that

$$\left(\mathbb{E}\left\|\sum_{j=1}^n g_j(\omega)x_j\right\|^2\right)^{1/2} \leq T\left(\sum_{j=1}^n \|x_j\|^p\right)^{1/p}$$

resp.

$$\left(\sum_{j=1}^n \|x_j\|^q\right)^{1/q} \leq C\left(\mathbb{E}\left\|\sum_{j=1}^n g_j(\omega)x_j\right\|^2\right)^{1/2}$$

for all $x_1, ..., x_n \in X$.

The following two statements assert that "Rademacher" (co)type and gaussian (co)type are somewhat comparable with each other.

Lemma 5.15. *If* $1 \leq p \leq 2 \leq q < \infty$, *then*

$$T_p(X,n) \leq \sqrt{\pi/2}\alpha_p(X,n) \leq \sqrt{\pi/2}K_pT_p(X,n) \quad \textit{and}$$
$$\beta_q(X,n) \leq \sqrt{\pi/2}C_q(X,n)\,.$$

Proof. See [138, p. 53-54]. □

The following gives the missing bound for the cotypes when $X = L^q$.

Proposition 5.16. *For all* $C < \infty$ *and* $q \in [2,\infty)$ *there is a constant* $K = K(C,q)$ *such that if* $\beta_q(X) \leq C$ *(the gaussian cotype* q *constant) then for all* n *and* $x_1, ..., x_n \in X$, *one has*

$$\|\sum_{j=1}^n g_jx_j\|_{L^q(X)} \leq K\|\sum_{j=1}^n \varepsilon_jx_j\|_{L^q(X)}$$

where g_j *are independent, symmetric,* L^2*-normalized, random Gaussians, and* ε_j *Rademacher signs. In particular,* $C_q(X) \leq K\beta_q(X)$.

Proof. See [138, Appendix II, Theorem 1]. □

Definition 5.17. Let $1 \leq n \leq m$, X be a normed space, and

$$\mathrm{Rad}_n X := \{\sum_{j=1}^n r_j(t)x_j : x_j \in X,\, j = 1, ..., n\}$$

denote the subspace of $L^2(X, \{-1,1\}^m)$ that is spanned by the first n Rademacher functions. If $f = \sum_{A\subseteq\{1,\dots,m\}} w_A \cdot x_A \in L^2(X, \{-1,1\}^m)$ (where $\{w_\alpha\}_{\alpha\in A}$ is any orthonormal basis in $L^2(\{-1,1\}^m)$ equipped with counting measure and $\{x_\alpha\}_{\alpha\in A} \in X$ are coefficients), then

$$\mathrm{Rad}_n f := \sum_{j=1}^{n} r_j \cdot x_{\{j\}}\,.$$

Lemma 5.18. *Let X be a normed space, $n \in \mathbb{N}$, and $1 < p \le 2$. Then*

$$C_{p'}(X,n) \le T_p(X^*,n) \le \|\mathrm{Rad}_n\| C_{p'}(X,n)\,.$$

In particular, if X^* has type p then X has cotype p'. Conversely, if $\|Rad_n\| < \infty$ and X has cotype p', then X^* has type p.

Proof. See [138, Lemma 9.10 and Corollary 9.11]. □

We now state a theorem estimating $\|\mathrm{Rad}_n\|$ for a general finite-dimensional space.

Theorem 5.19. *Let X be a finite-dimensional normed space of dimension k. Then, for all n, m, one has*

$$\|\mathrm{Rad}_n\|_{L^2(\{-1,1\}^m,X)} \le (e+1)\log(1 + d(X, \ell_k^2))\,. \tag{5.18}$$

In particular, there exists a universal constant $K > 0$ such that

$$\|\mathrm{Rad}_n\|_{L^2(\{-1,1\}^m,X)} \le K \cdot \log k\,. \tag{5.19}$$

Moreover, if $X \subseteq L^1(0,1)$, then

$$\|\mathrm{Rad}_n\|_{L^2(\{-1,1\}^m,X)} \le K \cdot (\log k)^{1/2}\,. \tag{5.20}$$

Proof. See [138, Theorem 14.5] and [24, p. 94]. □

5.3. **Tails of sub-gaussian distributed random variables.**

Definition 5.20 (Orlicz function[4]). An *Orlicz function* is a convex increasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\psi(0) \in [0,1)$. For a random variable X we define its *Orlicz norm* by

$$\|X\|_\psi := \inf\{c > 0 : \mathbb{P}[\psi(|X|/c)] \le 1\}$$

with the understanding that $\|X\|_\psi = \infty$ if the infimum runs over an empty set. The *Orlicz space* $L^\psi = L^\psi(\Omega, \Sigma, \mathbb{P})$ consists of all random variables X on the probability space $(\Omega, \Sigma, \mathbb{P})$ with finite Orlicz norm, i.e., $L^\psi = \{X : \|X\|_\psi < \infty\}$.

Example 5.21.

- For $q \ge 1$, the function $\psi_q(x) = \exp(x^q) - 1$ is an Orlicz function with $\|X\|_\psi < \infty$ if and only if there is $K_1 > 0$ such that $X - \mathbb{P}X$ satisfies $\mathbb{P}\{|X| > t\} \le 2\exp(-t^q/K_1^q)$ for all $t > 0$. (If $q = 2$, we will say that X is sub-gaussian.)
- For $p \in [1,\infty])$ the function $\psi(x) = x^p$ is Orlicz.
- We have the hierarchy $L^\infty \subseteq L^{\psi_2} \subseteq L^p$ for all $p \in [1,\infty)$. (The first inclusion is a consequence of (2) in Proposition 5.23 and the second inclusion follows from the obvious bound $\|X\|_{\psi_2} \lesssim \|X\|_\infty$.)

Remark 5.22. The bound $\|X\|_\psi \le \sigma$ immediately gives the tail bound

$$\mathbb{P}(|X| > t) \le \frac{\mathbb{P}\psi(|X|/\sigma)}{\psi(t/\sigma)} \le \frac{1}{\psi(t/\sigma)}\,, \quad t > 0\,.$$

[4]See [194, Section 2.7].

Proposition 5.23 (Sub-gaussian properties). *Let X be a random variable. Then the following are equivalent.*

1. *The tails of X satisfy $\mathbb{P}\{|X| > t\} \le 2\exp(-t^2/K_1^2)$ for all $t > 0$.*
2. *$\|X\|_{L^p} := (\mathbb{E}|X|^p)^{1/p} \le K_2\sqrt{p}$ for all $p \ge 1$.*
3. *$\mathbb{E}\exp(\lambda^2 X^2) \le \exp(K_3^2\lambda^2)$ for all $\lambda \in \mathbb{R}$ such that $|\lambda| < 1/K_3$.*
4. *There is $K_4 > 0$ such that $\mathbb{E}(\exp(X^2/K_4^2)) \le 2$. (This is called the ψ_2 condition.)*

Moreover, if $\mathbb{E}X = 0$ then (1)-(4) are also equivalent to

5. *$\mathbb{E}\exp(\lambda X) \le \exp(K_5^2\lambda^2)$ for all $\lambda \in \mathbb{R}$.*

The parameters K_i appearing in the statements differ from each other by at most an absolute factor.

Proof. See Vershynin [194, Proposition 2.5.2]. □

Definition 5.24. A random variable X that satisfies one of the equivalent properties in Proposition 5.23 is called a *sub-gaussian random variable*. The *sub-gaussian norm* of X, denoted by $\|X\|_{\psi_2}$ is defined to be the smallest K_4 in the fourth property in Proposition 5.23, i.e.,

$$\|X\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}(\exp(X^2/t^2)) \le 2\}. \tag{5.21}$$

Thus, if X is sub-gaussian, then, e.g.,

$$\mathbb{P}\{|X| > t\} \le 2\exp(-ct^2/\|X\|_{\psi_2}^2).$$

Moreover it is clear that when X is sub-gaussian, then so is $X - \mathbb{E}X$ with $\|X - \mathbb{E}X\|_{\psi_2} \lesssim \|X\|_{\psi_2}$ (by Jensen, but see also [194, Lemma 2.6.8]).

Example 5.25.

- Random gaussians $X \sim N(0, \sigma^2)$ with variance σ^2 are sub-gaussian with $\|X\|_{\psi_2} \lesssim \sigma$.
- Rademacher signs are sub-gaussian with $\|X\|_{\psi_2} = 1/\sqrt{\log 2}$ since $|X| = 1$.

See also Vershynin [194, Example 2.5.8].

The following lemma is crucial for the proof of Theorem 5.2.

Lemma 5.26 (Maximum of sub-gaussians). *Let $(X_j)_{j\in\mathbb{N}}$ be a sequence of sub-gaussian random variables which are not necessarily independent. Then we have for any $N \ge 2$ that*

$$\mathbb{E}\max_{1\le j\le N}|X_j| \lesssim \max_{1\le j\le N}\|X_j\|_{\psi_2}\sqrt{\log N}.$$

Remarks 5.27. (1) In some sense this lemma can be seen as a substitute for the dual of the missing $p = \infty$-Kintchine inequality

$$\int_{\mathbb{R}^d}\mathbb{E}|\sum_j \varepsilon_j f_j(x)|^p\,dx \sim \|(\sum_j |f_j|^2)^{1/2}\|_{L^p}^p. \tag{5.22}$$

(2) This estimate is sharp as can be seen by taking $X_1, ..., X_N$ to be N independent $N(0,1)$ normal distributed variables. Then $\mathbb{E}\max_{1\le j\le N} X_j \gtrsim \sqrt{\log N}$ (cf. [194, Exercise 2.5.11]). A complete proof can be found in [143, Theorem 3].

Proof. See Vershynin [194, Exercise 2.5.10] or Theorem 1.14 in MIT notes `https://ocw.mit.edu/courses/mathematics/18-s997-high-dimensional-statistics-spring-2015/lecture-notes/MIT18_S997S15_Chapter1.pdf`. For a complete proof, see Artstein–Giannopoulos–Milman [2, Proposition 3.5.8].

We give the proof under the additional assumption $\mathbb{E}X_j = 0$. First note that

$$\max_{1\le j\le N}|X_j| = \max_{1\le j\le 2N} X_j, \quad X_{N+\ell} := -X_\ell \text{ for } \ell = 1, ..., N,$$

so it suffices to prove an estimate for $\mathbb{E}\max_{1\leq j\leq N} X_j$. Abbreviate $K := \max_{1\leq j\leq N}\|X_j\|_{\psi_2}$. Then, we have by Jensen for any $s>0$,

$$\begin{aligned}\mathbb{E}\max_{1\leq j\leq N} X_j &= \frac{1}{s}\mathbb{E}[\log \mathrm{e}^{s\max X_j}] \leq \frac{1}{s}\log\left(\mathbb{E}[\mathrm{e}^{s\max X_j}]\right) = \frac{1}{s}\log\left(\mathbb{E}[\max_{1\leq j\leq N}\mathrm{e}^{sX_j}]\right)\\ &\leq \frac{1}{s}\log\left(\sum_{j=1}^N \mathbb{E}[\mathrm{e}^{sX_j}]\right) \leq \frac{1}{s}\log\left(\sum_{j=1}^N \mathrm{e}^{K^2s^2}\right) = \frac{\log(N)}{s} + sK^2\,.\end{aligned}$$

Optimizing the right side over $s>0$ gives $s=\sqrt{2\log(N)/K^2}$, i.e., the desired estimate. □

We will now state some tail bounds.

Lemma 5.28 (Sums of independent sub-gaussians)**.** *Let $X_1,...,X_N$ be independent, mean-zero, sub-gaussian random variables. Then $\sum_{n=1}^N X_n$ is sub-gaussian as well with*

$$\|\sum_{n=1}^N X_n\|_{\psi_2}^2 \lesssim \sum_{n=1}^N \|X_n\|_{\psi_2}^2\,.$$

Proof. See [194, Proposition 2.6.1]. □

This and Lemma 5.26 allow us to obtain the following corollary that will be crucial for the proof of Theorem 5.2. Compare also with [28, (4.1), (4.14)] and [20, (3.12)] where it is referred to as "Dudley's L^{ψ_2}-estimate".

Corollary 5.29. *Let $\{\omega_n\}_{n\in\mathbb{N}}$ be independent sub-gaussian random variables and $\mathcal{E}$ be a separable, (possibly infinite-dimensional) vector space over $\mathbb{C}$ with cardinality $|\mathcal{E}|$. Then*

$$\mathbb{E}\left(\sup_{\xi=(\xi_n)_{n\in\mathbb{N}}\in\mathcal{E}}|\sum_n \omega_n\xi_n|\right) \lesssim \sqrt{\log|\mathcal{E}|}\cdot \sup_{\xi=(\xi_n)_{n\in\mathbb{N}}\in\mathcal{E}}(\sum_n|\xi_n|^2)^{1/2}\,.$$

Proof. Identify X_j with $\sum_n \omega_n\xi_n^{(j)}$ where $(\xi_n^{(j)})_{n\in\mathbb{N}}$ denote the elements of the vector $\xi^{(j)}\in\mathcal{E}$ that we use to identify X_j. By Lemmas 5.26 and 5.28, one has

$$\mathbb{E}\left(\sup_{\xi\in\mathcal{E}}|\sum_n\omega_n\xi_n|\right) \leq \sqrt{\log|\mathcal{E}|}\sup_{\xi\in\mathcal{E}}\|\sum_n\omega_n\xi_n\|_{\psi_2} \leq \sqrt{\log|\mathcal{E}|}\sup_{\xi\in\mathcal{E}}[\sum_n|\xi_n|^2\|\omega_n\|_{\psi_2}^2]^{1/2}\,.$$

This concludes the proof. □

The following is a simple tail bound that is useful for measuring exceptional sets.

Proposition 5.30 (Hoeffding inequality)**.** *Let $X_1,...,X_N$ be independent, mean-zero, sub-gaussian random variables and $a=(a_1,...,a_N)\in\mathbb{R}^N$. Then for every $t>0$ we have*

$$\mathbb{P}(|\sum_{j=1}^N a_jX_j|\geq t) \leq 2\exp(-\frac{ct^2}{\sup_j\|X_j\|_{\psi_2}^2\|a\|_2^2})$$

Proof. See [194, Theorem 2.6.3]. □

Definition 5.31. A random variable X that satisfies $\|X\|_{\psi_1}$ where $\psi_1(x)=\mathrm{e}^x-1$ is called a *sub-exponential random variable.*

Proposition 5.32. *Let X be a random variable. Then the following are equivalent.*

(1) *The tails of X satisfy $\mathbb{P}\{|X|>t\}\leq 2\exp(-t/K_1)$ for all $t>0$.*
(2) *$\|X\|_{L^p} := (\mathbb{E}|X|^p)^{1/p}\leq K_2 p$ for all $p\geq 1$.*
(3) *$\mathbb{E}\exp(\lambda|X|)\leq\exp(K_3\lambda)$ for all $\lambda\in\mathbb{R}$ such that $|\lambda|<1/K_3$.*

(4) There is $K_4 > 0$ such that $\mathbb{E}(\exp(|X|/K_4)) \le 2$. (This is called the ψ_2 condition.)

Moreover, if $\mathbb{E}X = 0$ then (1)-(4) are also equivalent to

(5) $\mathbb{E}\exp(\lambda X) \le \exp(K_5^2\lambda^2)$ for all $\lambda \in \mathbb{R}$ such that $|\lambda| \le 1/K_5$.

The parameters K_i appearing in the statements differ from each other by at most an absolute factor.

Proof. See [194, Proposition 2.7.1]. □

Lemma 5.33. *(1) Any sub-gaussian random variable is also sub-exponential.*

(2) A random variable X is sub-exponential if and only if X^2 is sub-gaussian and in this case $\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2$.

(3) Let X and Y be sub-gaussian random variables. Then $\|XY\|_{\psi_1} \le \|X\|_{\psi_2}\|Y\|_{\psi_2}$.

(4) If X is sub-exponential, then $\|X - \mathbb{E}X\|_{\psi_1} \lesssim \|X\|_{\psi_1}$.

Theorem 5.34 (Bernstein)**.** *Let $X_1, ..., X_N$ be independent, mean-zero, sub-exponential random variables, and $a = (a_1, ..., a_N) \in \mathbb{R}^N$. Then for every $t \ge 0$ we have*

$$\mathbb{P}\left(|\sum_{j=1}^N a_j X_j| \ge t\right) \le 2\exp\left(-c\min\{\frac{t^2}{\max_j \|X_j\|_{\psi_1}^2 \|a\|_2^2}, \frac{t}{\max_j \|X_j\|_{\psi_1}\|a\|_\infty}\}\right).$$

Proof. See [194, Theorem 2.8.2]. □

5.4. **Sub-gaussians, Sudakov, Dudley, and entropy once again.** The following bounds once more refer to geometry of Banach spaces (Proposition 5.6 and Theorem 5.7) now in a concrete probabilistic setting.

Theorem 5.35 (Sudakov's minorization)**.** *Let $(X_t)_{t\in T}$ be a Gaussian process indexed by a set T equipped with the pseudo-metric[5] d_X induced by X defined as[6]*

$$d_X(s,t) = \|X_s - X_t\|_{L^2} = \left(\mathbb{E}(X_t - X_s)^2\right)^{1/2}, \quad s,t \in T.$$

Then for each $\varepsilon > 0$, we have

$$\log N(T, d_X, \varepsilon) \lesssim \frac{\sup_{t\in T}|X_t|^2}{\varepsilon^2}.$$

Proof. See [129, Theorem 3.18]. □

Example 5.36. Consider Brownian motion where $X_t - X_s \sim N(0, t-s)$, i.e., the increments are independent and are distributed according to the Gaussian law $d\mu_{t-s}(x) = (t-s)^{-1/2}\exp(-|x|^2/(t-s))\,dx$. Then

$$d_X(t,s)^2 = \int x^2\, d\mu_{t-s}(x) = \int x^2 \frac{\mathrm{e}^{-|x|^2/(t-s)}}{\sqrt{t-s}}\, dx \sim t - s.$$

Definition 5.37 (Sub-gaussian increments)**.** Consider a random process $(X_t)_{t\in T}$ on a metric space (T, d). We say that this process has *sub-gaussian increments* if there exists $K > 0$ such that

$$\|X_t - X_s\|_{\psi_2} \le K d(t,s), \quad t, s \in T. \tag{5.23}$$

[5]That is, $d(t,s) = 0$ does not necessarily imply $t = s$.

[6]The pseudo-metric $d_X(s,t)$ is also called "increments of the random process $(X_t)_{t\in T}$".

Example 5.38. Let $(X_t)_{t\in T}$ be a Gaussian process on an abstract set T. Define a metric on T by

$$d(t,s) := \|X_t - X_s\|_{L^2}\,, \quad t,s\in T\,.$$

Then $(X_t)_{t\in T}$ is a process with sub-gaussian increments and K above is an absolute constant.

We now state Dudley's inequality which gives a bound on a general sub-gaussian random process $(X_t)_{t\in\mathbb{T}}$ in terms of the metric entropy $\log(N(T,d,\varepsilon))$ of T. Note that it almost complements Sudakov's bound in Theorem 5.35.

Theorem 5.39 (Dudley)**.** *Let $(X_t)_{t\in T}$ be a mean-zero random process on a metric space (T,d) with sub-gaussian increments as in* (5.23)*. Then*

$$\mathbb{E}\sup_{t\in T} X_t \lesssim K\int_0^\infty \sqrt{\log(N(T,d,\varepsilon))}\,d\varepsilon$$

and

$$\mathbb{E}\sup_{t\in T} X_t \lesssim K\sum_{k\in\mathbb{Z}} 2^{-k}\sqrt{\log(N(T,d,2^{-k}))}$$

Proof. See [129, Theorem 11.17] or [194, Theorems 8.1.3 and 8.1.4]. □

Theorem 5.40 (Fernique)**.**

Proof. See [129, Theorem 11.18] and Theorem 6.6 in `https://www.math.ucla.edu/~biskup/PIMS/PDFs/lecture6.pdf`. □

5.5. **Proof of Theorems 5.2 and 5.3.** As mentioned in the beginning of this section, one of the main tools will be the dual to Sudakov bound (Theorem 5.7), i.e.,

$$\log N(B^n,\|\cdot\|_X,t) \lesssim n\cdot(\frac{A_r}{t})^2$$

where $B^n = \{x\in\mathbb{R}^n : \|x\|_2\leq 1\}$ is the euclidean unit ball in $\mathbb{R}^n$, $\|\cdot\|_X$ denotes another norm on $\mathbb{R}^n$, and

$$A_r = \int_{S^{n-1}} \|x\|_X\,d\mu(x) \sim n^{-1/2}\int_\Omega \|\sum_{j=1}^n g_j(\omega)e_j\|_X\,d\rho(\omega)$$

denotes the average of $\|x\|_X$ on the euclidean unit sphere which can be expressed probabilistically using n independent random gaussians and any orthonormal basis $\{e_j\}_{j=1}^n$ of $\mathbb{R}^n$, cf. (5.10).

Proof of Theorem 5.2. We only focus on $\Gamma_{t_1} = \Gamma_{t_2} \equiv \Gamma$. The general case is proven analogously. Our task is to compute

$$\|F_S V_\ell F_S^*\|_{L^2(\Gamma),L^2(\Gamma)} = \sup_{\mu_1,\mu_2}\left|\sum_{|n|\sim 2^\ell} V_\omega(n)\hat{\mu}_1(n)\hat{\mu}_2(n)\right| = 2^{-\varepsilon\ell}\sup_{\mu_1,\mu_2}\left|\sum_{|n|\sim 2^\ell}\omega_n v_n\hat{\mu}_1(n)\hat{\mu}_2(n)\right| \tag{5.24}$$

where the supremum is taken over all $\mu_j\in L^2(\Gamma,d\sigma)$ with $\|d\mu/d\sigma\|_2\leq 1$ for $j=1,2$. The main idea is to find a (finite) covering $\mathcal{E}_t$ of the distorted euclidean ball $\{\hat{\mu}(n)|_{|n|\sim 2^\ell} : \|\mu\|_{L^2(\Gamma)}\leq 1\}$ with $\ell^\infty_{|n|\sim 2^\ell}$-balls of radius t and to expand $\hat{\mu}(n)|_{|n|\sim 2^\ell} = \sum_r \xi^{(r)}(n)$ for some $\xi^{(r)}\in\mathcal{F}_r\subseteq \mathcal{E}_{2^{-r-1}} - \mathcal{E}_{2^{-r}}$. The main task is to understand $\|\xi^{(r)}\|_p$ for $p\in\{6,\infty\}$ and the cardinality $|\mathcal{F}_r|\lesssim|\mathcal{E}_{2^{-r-1}}|\cdot|\mathcal{E}_{2^{-r}}|$. Of course, the latter quantity will be estimated by means of the dual to Sudakov estimate (Theorem 5.7).

To apply the dual to Sudakov bound, we construct the norm $\|\cdot\|_X$ on $\mathbb{R}^d$ as follows. Consider a linear operator

$$S = (S_1, ..., S_m) : \ell^2_d \to \ell^\infty_m$$

(where each S_j has d columns for any $j = 1, ..., m$) and define

$$\|\psi\|_X := \|S\psi\|_{\ell^\infty_m} . \tag{5.25}$$

Now change the perspective and observe that not only does $N(B^d, \|\cdot\|_X, t)$ equal the minimal number of t-balls in $\|S\cdot\|_{\ell^\infty_m}$ norm needed to cover $\{\psi \in B^d\}$ but also the minimal number of t-balls in $\|\cdot\|_{\ell^\infty_m}$ norm needed to cover the deformed euclidean unit ball $\{S\psi : \psi \in B^d\} \subseteq \ell^\infty_m$. We will now compute the average. Using Corollary 5.29, we obtain

$$\begin{aligned} A_r &\sim d^{-1/2} \int_\Omega d\mu \max_{1\le j\le m} |\sum_{n=1}^d S_{j,n} e_n g_n(\omega)| \\ &\lesssim d^{-1/2} (\log m)^{1/2} \max_{1\le j\le m} (\sum_{n=1}^d |S_{j,n} e_n|^2)^{1/2} \\ &= d^{-1/2} (\log m)^{1/2} \|S\|_{\ell^2_d \to \ell^\infty_m} \end{aligned}$$

Thus, the entropy number for $\{S\psi : \psi \in B^d\}$ is bounded by

$$\log N(B^d, \|\cdot\|_X, t) \lesssim (\log m) t^{-2} \|S\|^2_{\ell^2_d \to \ell^\infty_m} ,$$

which – and this is important – does not depend on the dimension d that we started with. But that means that we may cover any infinite-dimensional euclidean ball, such as L^2, with balls in a suitable L^∞ metric. In fact, we will now replace ℓ^2_d by our space of interest, namely $L^2(\Gamma, d\Sigma)$. The role of S will be played by the localized Fourier extension operator

$$\begin{aligned} S : L^2(\Gamma, d\Sigma) &\to \ell^\infty(\mathbb{Z}^2) \\ \mu &\mapsto \hat{\mu}\big|_{|n|\sim 2^\ell} \end{aligned}$$

for some $\ell \in \mathbb{N}$. (Recall that 2^ℓ was the localization in physical space where we splitted $V = \sum_\ell V_\ell$ with $V_\ell = V \mathbf{1}_{2^{\ell-1} \le |n| \le 2^\ell}$.) Indeed, by its very definition (or Riemann–Lebesgue), we have

$$\|S\|_{L^2(\Gamma, d\Sigma) \to \ell^\infty(\mathbb{Z}^2)} \le C$$

and, by Tomas–Stein,

$$\|S\|_{L^2(\Gamma, d\Sigma) \to \ell^6(\mathbb{Z}^2)} \le C .$$

Thus, by the penultimate estimate and the above discussion, we can cover the set $\{S\psi : \psi \in L^2(\Gamma, d\Sigma), \|\psi\| \le 1\}$ with $N(t)$ many t-balls in the ℓ^∞_m-norm where now $m \sim 2^\ell$. Put differently, there exists a set $\mathcal{E}_t \subseteq \ell^\infty_{|n|\sim 2^\ell}$ of cardinality $|\mathcal{E}_t|$ that satisfies[7]

$$\begin{aligned} \log |\mathcal{E}_t| &< C\ell t^{-2} \\ \max_{\mu \in L^2(\Gamma),\, \|d\mu/d\Sigma\|_2 \le 1} \min_{\xi \in \mathcal{E}_t} \|\hat{\mu} - \xi\|_{\ell^\infty_{|n|\sim 2^\ell}} &< t \\ \max_{\xi \in \mathcal{E}_t} \|\xi\|_6 &< C \end{aligned} \tag{5.26}$$

We now take t of the form 2^{-r} for $r \in \mathbb{N}$. Thus, there exists a subset

$$\mathcal{F}_r \subseteq \mathcal{E}_{2^{-r-1}} - \mathcal{E}_{2^{-r}} = \{\xi_{r+1} - \xi_r : \xi_{r+1} \in \mathcal{E}_{2^{-r-1}}, \xi_r \in \mathcal{E}_{2^{-r}}\} \tag{5.27}$$

---

[7] Recall Remark 5.8.

with the properties[8]

$$\|\xi\|_\infty < 2^{-r+1} \quad \text{and} \quad \|\xi\|_6 < C\,, \quad \xi \in \mathcal{F}_r \tag{5.28}$$

and for each $\mu \in L^2(\Gamma)$ with $\|d\mu/d\sigma\|_2 \le 1$ there is a representation

$$S\mu = \hat{\mu}(n)\big|_{|n|\sim 2^\ell} = \sum_r \xi^{(r)} \quad \text{for some } \xi^{(r)} \in \mathcal{F}_r\,. \tag{5.29}$$

Plugging this decomposition into (5.24), we obtain

$$\|F_S V_\ell F_S^*\|_{L^2(\Gamma),L^2(\Gamma)} \le 2^{-\varepsilon\ell} \sum_{r_1,r_2\in\mathbb{N}} \max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\,\xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left|\sum_{|n|\sim 2^\ell} \omega_n v_n \xi_n^{(1)}\xi_n^{(2)}\right|\,. \tag{5.30}$$

Now fix r_1, r_2 and take the ω-expectation. On the one hand, we have the simple deterministic bound

$$\max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\,\xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left|\sum_{|n|\sim 2^\ell} \omega_n v_n \xi_n^{(1)}\xi_n^{(2)}\right| \lesssim 2^{-r_1-r_2} \sum_{|n|\sim 2^\ell} |\omega_n||v_n| \lesssim 2^{-r_1-r_2}\cdot 2^{4\ell/3}\|v\|_3\,, \tag{5.31}$$

which already behaves quite well in r_1, r_2 but terribly in ℓ. We now derive a second bound. Since the $\{\omega_n\}$ are independent sub-gaussian random variables, we may apply Dudley's estimate (Corollary 5.29) and obtain

$$\begin{aligned}
&\mathbb{E}_\omega\left[\max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\,\xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left|\sum_{|n|\sim 2^\ell} \omega_n v_n \xi_n^{(1)}\xi_n^{(2)}\right|\right] \\
&\lesssim \left(\log|\mathcal{F}_{r_1}^{(1)}| + \log|\mathcal{F}_{r_2}^{(2)}|\right)^{1/2} \left[\max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\,\xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left(\sum_{|n|\sim 2^\ell} |v_n|^2|\xi_n^{(1)}|^2|\xi_n^{(2)}|^2\right)^{1/2}\right]\,.
\end{aligned} \tag{5.32}$$

To estimate $\log|\mathcal{F}_{r_j}^{(j)}|$, recall that (5.26) said $\log|\mathcal{E}_{2^{-r}}| \lesssim \ell 4^r$. Combining this with the trivial estimate $|\mathcal{F}_r| \lesssim |\mathcal{E}_{2^{-r-1}}|\cdot|\mathcal{E}_{2^{-r}}|$ gives

$$\log|\mathcal{F}_r| \lesssim \log|\mathcal{E}_{2^{-r-1}}| + \log|\mathcal{E}_{2^{-r}}| \lesssim \ell 4^r\,. \tag{5.33}$$

Next, we bound [...] on the right side of (5.32). Using Hölder and (5.28), the sum over $|n|\sim 2^\ell$ is bounded by

$$\left(\sum_{|n|\sim 2^\ell} |v_n|^2|\xi_n^{(1)}|^2|\xi_n^{(2)}|^2\right)^{1/2} \le \|v\|_3\|\xi^{(1)}\cdot\xi^{(2)}\|_6 \lesssim \|v\|_3 \min\{2^{-r_1}, 2^{-r_2}\}\,. \tag{5.34}$$

Combining (5.33) with (5.34), we obtain

$$\mathbb{E}_\omega\left[\max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\,\xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left|\sum_{|n|\sim 2^\ell} \omega_n v_n \xi_n^{(1)}\xi_n^{(2)}\right|\right] \lesssim \sqrt{\ell}(2^{r_1}+2^{r_2})\min\{2^{-r_1},2^{-r_2}\}\|v\|_3 \lesssim \sqrt{\ell}\|v\|_3\,.$$

[8] One may think of $\mathcal{F}_r$ being the set of differences $\xi - \xi'$ where ξ and ξ' belong to the same or "parental" 2^{-r} or $2r^{-r-1}$ cube in $\ell^\infty_{|n|\sim 2^\ell}$ metric. This is a consequence of the geometrical fact that dyadic cubes either contain each other or are disjoint.

This estimate alone would not be good enough to survive the r_1, r_2 summation. However, combining it with (5.31), we see that the ω-expectation of (5.30) is bounded from above by

$$\mathbb{E}_\omega \|F_S V_\ell F_S^*\|_{L^2(\Gamma),L^2(\Gamma)} \lesssim 2^{-\varepsilon\ell} \sum_{r_1,r_2\in\mathbb{N}} \min\{\sqrt{\ell}, 2^{-r_1-r_2}2^{4\ell/3}\}\|v\|_3 \lesssim \ell^{5/2}2^{-\varepsilon\ell}\|v\|_3 \lesssim 2^{-\varepsilon'\ell}\|v\|_3$$

for some $0 < \varepsilon' < \varepsilon$. This proves Theorem 5.2 for $V_\omega(n) = \omega(n)|n|^{-\varepsilon}v(n)$ and $v \in \ell^3(\mathbb{Z}^2)$.

If $V_\omega(n) = \omega(n)w(n)$ with $w \in \ell^{3-\varepsilon}(\mathbb{Z}^2)$, then the deterministic bound in (5.31) becomes

$$\max_{\xi^{(1)}\in\mathcal{F}_{r_1}^{(1)},\, \xi^{(2)}\in\mathcal{F}_{r_2}^{(2)}} \left| \sum_{|n|\sim 2^\ell} \omega_n w_n \xi_n^{(1)}\xi_n^{(2)} \right| \lesssim 2^{-r_1-r_2} \sum_{|n|\sim 2^\ell} |\omega_n||w_n| \lesssim 2^{-r_1-r_2}\cdot 2^{\frac{2(2-\varepsilon)}{3-\varepsilon}\ell}\|w\|_{3-\varepsilon}$$

which is – as expected – a slight improvement over (5.31) since $(2-\varepsilon)/(3-\varepsilon) < 2/3$. On the other hand, (5.34) (which came from the probabilistic estimate using Dudley's L^{ψ_2} inequality) is improved to

$$\left(\sum_{|n|\sim 2^\ell} |w_n|^2|\xi_n^{(1)}|^2|\xi_n^{(2)}|^2\right)^{1/2} \leq \|w\|_{3-\varepsilon}\|\xi^{(1)}\cdot\xi^{(2)}\|_{\frac{2(3-\varepsilon)}{1-\varepsilon}} \lesssim \|w\|_{3-\varepsilon}2^{-\tilde\varepsilon(r_1+r_2)}\cdot\min\{2^{-r_1},2^{-r_2}\}$$

for $\tilde\varepsilon = (p-6)/p > 0$ with $p = 2(3-\varepsilon)/(1-\varepsilon) > 6$ which follows from Hölder's inequality

$$\|\xi^{(j)}\|_p \leq \|\xi^{(j)}\|_6^{6/p}\|\xi^{(j)}\|_\infty^{\frac{p-6}{p}} \lesssim 2^{-\frac{p-6}{p}r_j}\,.$$

Combining these estimates as before then gives

$$\mathbb{E}_\omega\|F_SV_\ell F_S^*\|_{L^2(\Gamma),L^2(\Gamma)} \lesssim \sum_{r_1,r_2\in\mathbb{N}} \min\{2^{-\tilde\varepsilon(r_1+r_2)}\sqrt{\ell}, 2^{-r_1-r_2}2^{\frac{2(2-\varepsilon)}{3-\varepsilon}\ell}\}\|w\|_{3-\varepsilon} \lesssim 2^{-\varepsilon'\ell}\|w\|_{3-\varepsilon}\,.$$

This concludes the proof of Theorem 5.2. □

Proof of Theorem 5.3. Again, we only focus on $\Gamma_{t_1} = \Gamma_{t_2} \equiv \Gamma$. The general case is proven analogously.

We split again $V_\omega(x) = \sum_{\ell\geq 0} V_\ell(x)$ with $V_0(x) = V_\omega(x)\mathbf{1}_{|x|\leq 1}(x)$ and, for $\ell\in\mathbb{N}$, $V_\ell(x) = V_\omega(x)\mathbf{1}_{2^{\ell-1}\leq|x|\leq 2^\ell}(x)$, and estimate

$$\mathbb{E}_\omega\|F_S^*V_\omega F_S\| \leq \sum_{\ell\geq 0}\mathbb{E}_\omega\|F_S^*V_\ell F_S\|\,.$$

To estimate the right side, we again use the dual to Sudakov bound. The norm $\|\cdot\|_X$ is essentially the same as in the proof of Theorem 5.2 by taking $\|\psi\|_X := \|S\psi\|_{L^\infty_{|x|\sim 2^\ell}}$ where $S := F_S^* : L^2(\Gamma, d\Sigma) \to L^\infty$ acts as $F_S^*g(x) = \int_\Gamma e^{2\pi i x\cdot\xi}g(\xi)d\Sigma(\xi)\big|_{|x|\sim 2^\ell}$. The theorem is concluded by repeating all the other steps of the proof of Theorem 5.2. □

5.6. **Some applications of random Tomas–Stein Theorem 5.3.**

5.6.1. *Complex Lieb–Thirring.* As is customary, we invoke the Birman–Schwinger principle to derive bounds on the modulus of complex eigenvalues. This requires to estimate

$$\|(T(\xi)-z)^{-1/2}V(T(\xi)-z)^{-1/2}\|\,.$$

If $T(\xi)$ is homogeneous, which is, e.g., the case when $T(\xi) = \xi^2$, we can always rescale and assume $|z| = 1$. Splitting into the regions where $T(\xi) \lesseqgtr \tau$, where $\tau \gg |z|$, it suffices (by Cauchy–Schwarz) to estimate the right side of

$$\begin{aligned}
&\|(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\| \\
&\quad + 2\|(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T>\tau\}}\| \\
&\quad + \|(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T>\tau\}}V(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T>\tau\}}\| \\
&\leq \|(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\| \\
&\quad + 2\||V|^{1/4}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\| \cdot \||V|^{3/4}(T-z)^{-1/2}\mathbf{1}_{\{T>\tau\}}\| \\
&\quad + \||V|^{1/2}(T(\xi)-z)^{-1}\mathbf{1}_{\{T>\tau\}}|V|^{1/2}\|\,.
\end{aligned}$$

The second factor of the second summand as well as the third term on the right are harmless and can easily be controlled even in Schatten spaces using Kato–Seiler–Simon whenever $V \in L^q$. The first factor of the second summand is controlled by Tomas–Stein or Kenig–Ruiz–Sogge [120, 87] whenever $V \in L^{d+1}$. (Any other splitting of $V = \mathrm{sgn}(V)|V|^\alpha \cdot |V|^{1-\alpha}$ would work equally well whenever $V \in L^p$ for an appropriate $\alpha(p)$.)

For a cut-off χ in physical space, we automatically have

$$\chi_{\{|n|<N_1\}}(T(\xi)-z)^{-1}\mathbf{1}_{\{T<\tau\}}\chi_{\{|n|<N_2\}} = \chi_{\{|n|<N_1\}}\left((T(\xi)-z)^{-1}\mathbf{1}_{\{T<\tau\}} * \gamma_{1/N}\right)\chi_{\{|n|<N_2\}}$$

where $N > N_1 + N_2$ and $\gamma_{1/N}(\xi) = N^d\gamma(N\xi)$ and $\check{\gamma}$ is a smooth bump on $\mathbb{R}^d$ with $\check{\gamma}(x) = 1$ for $|x| < 1$.

The main goal is to control

$$\|(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V_\omega(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\|$$

by expressions like

$$\|[(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} * \gamma_{\delta_1}]V_\ell[(T(\xi)-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} * \gamma_{\delta_2}]\| \tag{5.35}$$

where $\delta_1, \delta_2 \sim 2^{-\ell}$. To simplify notation, introduce

$$C^{(\infty)}(\xi) := \frac{\mathbf{1}_{\{T(\xi)<\tau\}}}{(T(\xi)-z)^{1/2}}$$

and

$$\begin{aligned}
C^{(\delta)}(\xi) &:= \left(\frac{\mathbf{1}_{\{T(\xi)<\tau\}}}{(T(\xi)-z)^{1/2}} * \gamma_\delta\right)(\xi) = \int_{\mathbb{R}^d} d\eta \frac{\mathbf{1}_{\{T(\eta)<\tau\}}}{(T(\eta)-z)^{1/2}}\gamma_\delta(\eta-\xi) \\
&= \int_0^\tau dt\,\frac{1}{(t-z)^{1/2}}\int_{S_t} d\Sigma_{S_t}(\eta)\,\gamma_\delta(\eta-\xi)\,.
\end{aligned} \tag{5.36}$$

Thus, the convolution essentially smoothes $(t-z)^{-1/2}\mathbf{1}_{\{t<\tau\}}$ out on the scale δ, and so

$$\begin{aligned}
|C^{(\delta)}(\xi)| &\lesssim \left[|T(\xi)-z|^{1/2}+\delta\right]^{-1}\mathbf{1}_{\{T(\xi)<10\tau\}} + \langle|\xi|\rangle^{-\tilde{N}}\mathbf{1}_{\{T(\xi)>10\tau\}} \\
&\lesssim \left[|T(\xi)-z|^{1/2}+\delta\right]^{-1}\mathbf{1}_{\{T(\xi)<10\tau\}} + T(\xi)^{-N}\mathbf{1}_{\{T(\xi)>10\tau\}} \equiv \widetilde{C^{(\delta)}}(\xi)\,.
\end{aligned} \tag{5.37}$$

for any $\tilde{N}, N \in \mathbb{N}$ by ellipticity of $T(\xi)$.

Comparing with (5.35) we are concerned with the elementary operators

$$C^{(\delta_2)}V_\ell C^{(\delta_1)}\,. \tag{5.38}$$

Let $\max\{\sup_n |v_n||n|^\rho, \|v\|_{\ell^{d+1}}\} < \kappa$. We shall first prove

Proposition 5.41. *The random Tomas–Stein Theorem 5.3, i.e.,*

$$\mathbb{E}_\omega \|F_{S_{t_2}} V_\ell F_{S_{t_1}}^* \|_{L^2(\Gamma_{t_1}, d\Sigma_{t_1}), L^2(\Gamma_{t_2}, d\Sigma_{t_2})} \lesssim_{|V|} 2^{-c_{|V|}\ell} ,$$

implies

$$\mathbb{E}_\omega \|C^{(\delta_1)} V_\ell C^{(\delta_2)}\|_{L^2(\mathbb{R}^d), L^2(\mathbb{R}^d)} \lesssim_{|V|} (\kappa 2^{-\ell})^{c_{|V|}} \left(\log \frac{1}{\delta_1} + \log \frac{1}{\delta_2}\right)^A \tag{5.39}$$

and

$$\mathbb{E}_\omega \|F_S V_\ell C^{(\delta)}\|_{L^2(\mathbb{R}^d), L^2(\Gamma, d\Sigma)} \lesssim_{|V|} (\kappa 2^{-\ell})^{c_{|V|}} (\log \frac{1}{\delta})^A \tag{5.40}$$

for $|z| = 1$.

Remarks 5.42. (1) Observe that $\delta_j \sim 2^{-\ell}$ is reasonable and leads to an expression that is still summable in ℓ.

(2) Recall [28, Lemmas 3.18 and 3.48], cf. [20, Lemmas 2.7 and 2.10] which are very similar to (5.39) and (5.40). There C^δ was either $C_1^{(\delta)}(\xi) = |(T(\xi) - z)^{-1} * \gamma_\delta|^{1/2}$ or $C_2^{(\delta)}(\xi) = \left((T(\xi) - z)^{-1} * \gamma_\delta\right)^{1/2}$. Here, we defined $C^\delta(\xi) = (T(\xi) - z)^{-1/2} * \gamma_\delta$.

Proof. By Plancherel and (5.37) it suffices to consider

$$\mathbb{E}_\omega \|F_S V_\ell \widetilde{C^{(\delta)}}\|_{L^2(\mathbb{R}^d), L^2(\Gamma, d\Sigma)} . \tag{5.41}$$

instead of the left side of (5.40). Neglecting the ω-expectation for a moment, the left side of (5.41) equals

$$\|F_S V_\ell \widetilde{C^{(\delta)}}\|_{L^2(\mathbb{R}^d), L^2(\Gamma, d\Sigma)} = \sup_{\mu \in L^2(S), g \in L^2(\mathbb{R}^d)} |(\mu, F_S V_\ell \widetilde{C_2^{(\delta)}} g)|$$

where the supremum is taken over all measures $\mu \ll \Sigma$ supported on Γ with $\|d\mu/d\Sigma\|_{L^2(\Gamma, d\Sigma)} \leq 1$ and all $g \in L^2(\mathbb{R}^d)$ with $\|g\|_{L^2(\mathbb{R}^d)} \leq 1\|$. Making use of the definition of $\widetilde{C^{(\delta)}}$ in (5.37), the spectral theorem (or the coarea formula) allows us to write

$$\begin{aligned}
\widetilde{C_2^{(\delta)}} g(x) &= \int_0^\infty dt \int_{S_t} d\Sigma_t(\xi)\, \mathrm{e}^{2\pi i x \cdot \xi} \widetilde{C^{(\delta)}}(\xi) (F_{S_t} g)(\xi) \\
&= \int_0^\infty dt \left(\left[|t - z|^{1/2} + \delta \right]^{-1} \mathbf{1}_{\{t < 10\tau\}} + t^{-N} \mathbf{1}_{\{t > 10\tau\}} \right) (F_{S_t}^* F_{S_t} g)(x) \\
&\equiv \int_0^\infty dt\, \widetilde{C^{(\delta)}}(t) (F_{S_t}^* F_{S_t} g)(x) .
\end{aligned}$$

Using the random Tomas–Stein estimate and (5.37), we obtain

$$\begin{aligned}
&\sup_{\mu, g} |(\mu, F_S V_\ell \widetilde{C_2^{(\delta)}} g)| \\
&= \sup_{\mu, g} | \int_0^\infty dt\, \widetilde{C^{(\delta)}}(t) \int_{\mathbb{R}^d} dx\, (F_S^* \mu)(x) V_\ell(x) (F_{S_t}^* F_{S_t} g)(x)| \\
&= \sup_{\mu, g} | \int_0^\infty dt\, \widetilde{C^{(\delta)}}(t) (\mu, F_S V_\ell F_{S_t}^* F_{S_t} g)| \\
&\lesssim 2^{-c_{|V|}\ell} \sup_{\|g\|_{L^2(\mathbb{R}^d)} \leq 1} \int_0^\infty dt \left(\left[|t - z|^{1/2} + \delta \right]^{-1} \mathbf{1}_{\{t < 10\tau\}} + t^{-N} \mathbf{1}_{\{t > 10\tau\}} \right) \|F_{S_t} g\|_{L^2(\Gamma_t, d\Sigma_t)} \\
&\lesssim_N 2^{-c_{|V|}\ell} [\log(|z|^{1/2}/\delta)]^{1/2} \|g\|_{L^2(\mathbb{R}^d)} = 2^{-c_{|V|}\ell} [\log(1/\delta)]^{1/2}
\end{aligned}$$

by Cauchy–Schwarz. Here, we used

$$\int_0^{10|z|} \frac{dt}{|t-z|+\delta^2} \lesssim \log\left(1+\frac{|z|}{\delta^2}\right) \sim \log\frac{1}{\delta}\,.$$

The proof of (5.39) is analogous and omitted. □

We show the following Birman–Schwinger bound.

Proposition 5.43. *Assume $T(\xi)$ is homogeneous, $|z|=1$, and let V_ω be so such that Theorem 5.3 is applicable. Let $\kappa = \min\{\sup_{n\in\mathbb{Z}^d} |n|^\rho |v(n)|, \|v\|_{d+1}\}$, then*

$$\mathbb{E}_\omega \|(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\| \leq C_{\rho,\kappa}\,. \tag{5.42}$$

Proof. We begin with a warm-up and consider $V_\omega(n) = \omega_n v_n$ with $\sup_n |v_n||n|^\rho = \kappa < \infty$, some $\rho > 1/2$, and consider only the truncated operator

$$\mathbf{1}_{|x|<N}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}$$

Since $(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}$ is normal, we have

$$\begin{aligned}
&\|\mathbf{1}_{|x|<N}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}\|^2 \\
&\quad = \|\mathbf{1}_{|x|<N}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\omega (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}\| \\
&\quad \leq \kappa^{2/3} \sum_{\ell,\ell'\geq 0} \|\mathbf{1}_{|x|<N}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} V_\ell (T-z)^{-1}\mathbf{1}_{\{T<\tau\}} V_{\ell'} (T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}\|\,.
\end{aligned}$$

Recall [28, Lemma 2.3]

$$\|V_\ell (T-z)^{-1} V_{\ell'}\| \lesssim 2^{-\frac{1}{2}(\rho-\frac{1}{2})(\ell+\ell')}\,.$$

Thus, we are left to bound

$$\|\mathbf{1}_{|x|\sim 2^\ell}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}} f\|,$$

for $f \in L^2$ with $\operatorname{supp} f \subseteq B_0(N)$. Equivalently, by TT^*, we bound

$$\langle f, \mathbf{1}_{|x|\sim 2^\ell}(T-\overline{z})^{-1/2}(T-z)^{-1/2}\mathbf{1}_{T<\tau}\mathbf{1}_{|x|\sim 2^\ell} f\rangle\,.$$

By the spatial cutoff, we can replace the Fourier multiplier in the middle by

$$F(\xi) := \Big((T(\cdot)-\overline{z})^{-1/2}(T(\cdot)-z)^{-1/2}\mathbf{1}_{T<\tau}\Big) * \gamma_{2^{1-\ell}}(\xi)$$

where

$$\gamma_\delta(\eta) = \delta^{-d}\gamma(\eta/\delta)\,.$$

Obviously, for any $M, \tilde{M} > 0$,

$$\begin{aligned}
|F(\xi)| &\lesssim_M ((T(\xi)^2 + |z|^2 - 2T(\xi)\mathrm{Re}(z))^{1/2} + 2^{-\ell})^{-1} + \langle\xi\rangle^{-\tilde{M}}\mathbf{1}_{T>10\tau} \\
&\lesssim_M ((T(\xi)^2 + |z|^2 - 2T(\xi)\mathrm{Re}(z))^{1/2} + 2^{-\ell})^{-1} + T(\xi)^{-M}\mathbf{1}_{T>10\tau}\,.
\end{aligned}$$

Suppose $\mathrm{supp} f \subseteq B_0(N)$, some $N > 0$. Then, by arguments similar to those in (5.37), we obtain (for $\tau \gg |z|$, say $\tau = 2|z|$, and assuming $|z| = 1$ with $\mathrm{Im}(z) \ll 1$)

$$\begin{aligned}
&\langle f, \mathbf{1}_{|x|\sim 2^\ell}(T-\bar{z})^{-1/2}(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N\wedge 2^\ell} f\rangle \\
&\lesssim_M \int dt \left[\frac{\mathbf{1}_{0<t<2|z|}}{(t^2+|z|^2-2t\mathrm{Re}(z))^{1/2}+2^{-\ell}+\frac{1}{N}} + t^{-M}\mathbf{1}_{t>2|z|}\right] \|F_{S_t}\mathbf{1}_{|x|<N\wedge 2^\ell} f\|^2 \\
&\lesssim_M \min_{\varepsilon>0} \varepsilon^{-1}(N\wedge 2^\ell)^{1+\varepsilon}\left[\int_0^{2|z|}\frac{dt}{|t-1|+|\mathrm{Im}(z)|+2^{-\ell}+1/N} + |z|^{1-M}\right] \qquad (5.43)\\
&\lesssim \log(N\wedge 2^\ell)(N\wedge 2^\ell)\log\left(\frac{1}{|\mathrm{Im}(z)|+2^{-\ell}+1/N}\right) \\
&\lesssim \ell^2 (N\wedge 2^\ell)\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbb{E}_\omega \|(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V_\omega(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\mathbf{1}_{|x|<N}\|^2 \\
&\lesssim \kappa^{2/3}\sum_{\ell,\ell'\geq 0} 2^{-\frac{1}{2}(\rho-\frac{1}{2})(\ell+\ell')}\cdot \ell\ell'(N\wedge 2^\ell)^{1/2}(N\wedge 2^{\ell'})^{1/2}
\end{aligned}$$

If $\rho > 1$, we would have uniform convergence in N. But that just corresponds just to the non-deterministic situation. □

Alternative proof of Proposition 5.43. We now consider the full operator and get rid of the additional spatial cut-off $|x| < N$ above.

We mimick arguments of Bourgain [28, pp. 13]. There, $Sf := \mathbf{1}_{|x|\sim 2^\ell}\mathcal{F}^{-1}C^{(k)}f$ with $C^{(k)} = |(T(\xi)-z)^{-1} * \gamma_{2^{-k}}|^{1/2}$. Here

$$Sf := \mathbf{1}_{|x|\sim 2^\ell}(T(\xi)-z)^{-1/2}\,.$$

The estimate follows once we verify [28, (4.5)-(4.6)]

Recalling the arguments from the previous proof (see (5.43)), we have

$$\|Sf\|_2 \lesssim \ell 2^{\ell/2}\|f\|_2$$

which is the analog of [28, (4.6)] Next, we show

$$\|Sf\|_\infty \lesssim \sqrt{\ell}\|f\|_2\,.$$

Again, we argue by SS^* which is given by

$$SS^* = \mathbf{1}_{|x|\sim 2^\ell}(T(\xi)-z)^{-1/2}(T(\xi)-\bar{z})^{-1/2}\mathbf{1}_{|x|\sim 2^\ell}\,.$$

and we hope to show $SS^* : L^1 \to L^\infty$. Let $f, g \in L^1$, then (for $|z| = 1$ and $\mathrm{Im}(z) \ll 1$)

$$\begin{aligned}
|(g, SS^* f)| &\lesssim \int dt \left[\frac{\mathbf{1}_{0<t<\tau}}{(t^2+|z|^2-2t\mathrm{Re}(z))^{1/2}+2^{-\ell}} + t^{-M}\mathbf{1}_{t>\tau}\right] |(g, F_{S_t}^* F_{S_t} f)| \\
&\lesssim \|g\|_1\|f\|_1 \log 2^\ell = \ell\|g\|_1\|f\|_1
\end{aligned}$$

since $F_{S_t} : L^1 \to L^2$ locally uniformly in t. This is the analog of [28, (4.5)].

This gives (cf. [28, (4.15)])

$$\begin{aligned}
&\mathbb{E}_\omega\|(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}V_\ell(T-z)^{-1/2}\mathbf{1}_{\{T<\tau\}}\| \\
&\lesssim \kappa 2^{-\rho\ell}(\ell+\log\frac{1}{\kappa})^2\cdot \ell^2\cdot 2^{\ell/2} \lesssim \kappa(\ell+\log(1+\frac{1}{\kappa}))^4 2^{-(\rho-\frac{1}{2})\ell}\,. \qquad (5.44)
\end{aligned}$$

□

As was remarked in [20, p. 74] this assertion also follows from Theorem 5.3.

5.6.2. *Weak coupling limit.* The following result is desirable but not yet proved.

Theorem 5.44 (Weak coupling limit). *Consider $H_\lambda^{(\omega)} := |\Delta+1| - \lambda V_\omega$ in $L^2(\mathbb{R}^d)$ with $V_\omega(x) = \sum_{n\in\mathbb{Z}^d} v_n\omega_n \mathbf{1}_{Q_n}(x)$ and $(v_n)_{n\in\mathbb{Z}^d} \subseteq \ell^{d+1-\varepsilon}(\mathbb{Z}^d)$. Let $\Gamma_t := \{\xi\in\mathbb{R}^d:\, |\xi^2-1| = t\}$ and $\Gamma := \Gamma_0$. Then for any eigenvalue a_Γ^j of $\mathcal{V}_\Gamma := F_S^* V_\omega F_S$ in $L^2(\Gamma, d\Sigma)$ there is an eigenvalue $e_j(\lambda)$ of H_λ satisfying the weak-coupling limit*

$$e_j(\lambda) = \exp\left(-\frac{1}{\lambda a_S^j}(1+o(1))\right), \quad \lambda\to 0$$

except for ω in a set of measure at most

$$\sum_{\ell\geq 0} \exp\left(-\frac{2^{\varepsilon'\ell}}{\|v\|_{d+1-\varepsilon}}\right) \lesssim_{\varepsilon'} \mathrm{e}^{-c_{\varepsilon'}/\|v\|_{d+1-\varepsilon}}\,.$$

Remark 5.45. Is a generalization of Theorems 5.2 and 5.3 to Schatten ideals possible? See Frank–Sabin [87, Theorem 2].

Proof of Theorem 5.44. By Theorem 5.3, the operator $\mathcal{V}_S = F_S V_\omega F_S^*$ is well-defined and isospectral (up to zero-eigenvalues that we neglect) to $|V_\omega|^{1/2} F_S^* F_S V_\omega^{1/2}$. `[Moreover, it is compact by Remark 5.45.]` Let $BS(e) = |V_\omega|^{1/2}(T+e)^{-1}V_\omega^{1/2}$ which is compact by Kato–Seiler–Simon. By the proof of [56, Theorem 1.1] it suffices to show that

$$|\lambda_j(BS(e)) - \lambda_j(\log(1/e)\mathcal{V}_S)| = o(\log(1/e))$$

where $\lambda_j(A)$ denotes the j-th non-zero eigenvalue of A. As usual, the high-energy piece

$$\|BS^{\mathrm{high}}(e)\| = \||V_\omega|^{1/2}(T+e)^{-1}\mathbf{1}_{[\tau,\infty)}(T)V_\omega^{1/2}\| \lesssim_\tau 1$$

is harmless, so we only focus on $BS^{\mathrm{low}}(e) = BS(e) - BS^{\mathrm{high}}(e)$. By isospectrality and Weyl's perturbation theorem [10, Corollary III.26], we have

$$\begin{aligned}&\sup_j |\lambda_j((T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)V_\omega(T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)) - \lambda_j(\log(1/e)F_S V_\omega F_S^*)|\\ &\quad\leq \|(T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)V_\omega(T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)) - \log(1/e)F_S V_\omega F_S^*\|\,.\end{aligned}$$

Decomposing $V_\omega = \sum_{\ell\geq 0} V_\ell$, it then suffices to show

$$\sum_{\ell\geq 0} \|(T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)V_\ell(T+e)^{-1/2}\mathbf{1}_{[0,\tau]}(T)) - \log(1/e)F_S V_\ell F_S^*\| = o(\log(1/e))\,.$$

`[The following arguments are not working out yet.]` We would like to bring the "regular part of the Birman–Schwinger operator"

$$BS_{\mathrm{reg}}^{\mathrm{low}}(e) := |V_\omega|^{1/2}[(T+e)^{-1}\mathbf{1}_{[0,\tau]}(T) - \log(1/e)F_S^*F_S]V_\omega^{1/2} \equiv |V_\omega|^{1/2}T_\tau V_\omega^{1/2}$$

into the play and show that it satisfies

$$\mathbb{E}_\omega \max\{z\in\mathbb{C}:\, z\in\mathrm{spec}(BS_{\mathrm{reg}}^{\mathrm{low}}(e))\} = o(\log(1/e))\,.$$

Since $BS_{\mathrm{reg}}^{\mathrm{low}}(e)$ is isospectral to $|T_\tau|^{1/2}V_\omega T_\tau^{1/2} = \sum_{\ell\geq 0}|T_\tau|^{1/2}V_\ell T_\tau^{1/2}$, we merely need to show (by a second application of isospectrality and making use of the fact that the spectral radius is always bounded from above by the operator norm),

$$\begin{aligned}&\mathbb{E}_\omega \max\{z \in \mathbb{C} :\ z \in \operatorname{spec}(\sum_{\ell \geq 0} |T_\tau|^{1/2} V_\ell T_\tau^{1/2})\} \\ &= \mathbb{E}_\omega \max\{z \in \mathbb{C} :\ z \in \operatorname{spec}(\sum_{\ell \geq 0} |V_\ell|^{1/2} T_\tau V_\ell^{1/2})\} \\ &\leq \sum_{\ell \geq 0} \mathbb{E}_\omega \| |V_\ell|^{1/2} T_\tau V_\ell^{1/2} \| = o(\log(1/e))\end{aligned}$$

However, this would follow from the spectral theorem and Hölder continuity of the non-endpoint random Tomas–Stein theorem, i.e.,

$$\mathbb{E}_\omega \sup_{t \in (0,\tau)} \| |V_\ell|^{1/2} (F_{S_t}^* F_{S_t} - \sqrt{1 \pm t} F_S^* F_S) V_\ell^{1/2} \|_{L^2(\mathbb{R}^d), L^2(\mathbb{R}^d)} \lesssim 2^{-c\ell} t^\alpha$$

for some $c, \alpha > 0$ depending solely on d and ε. But this is just a consequence of Theorem [abstractrandomts] which is, indeed, applicable due to [56, Proposition 4.4] which asserts the corresponding, non-deterministic Hölder continuity. □

6. Local restriction estimates

We follow Lecture 1 in Hickman–Vitturi [107] and strongly advise to consider Tao–Vargas–Vega [188] where the techniques that we are about to describe were first developed and systematically applied.

We now discuss the first tool which is used to prove the above restriction theorems. The key idea is to reduce the study of *global* restriction theorems (where the "physical" space variable is allowed to range over all $\mathbb{R}^d$) to *local* restriction theorems (where the physical space variable is constrained to lie in a ball). Our aim is then to prove estimates of the form

$$\|\hat{f}|_S\|_{L^q(S,d\sigma)} \leq A_{p,q,S,\alpha} R^\alpha \|f\|_{L^p(B(x_0,R))} \tag{6.1}$$

for any exponents p, q, $\alpha \geq 0$, and any radius $R \geq 1$ of any ball $B(x_0, R) = \{x \in \mathbb{R}^d : |x - x_0| \leq R\}$. We will denote the statement such that the above estimate holds for any test function f by $R_S(p \to q; \alpha)$. Note that the center x_0 of the ball is irrelevant since one can translate f by an arbitrary amount without affecting the *magnitude* of $\hat{f}$.

Obviously, we have $R_S(p \to q; \alpha_1) \Rightarrow R_S(p \to q; \alpha_2)$ if $\alpha_1 \leq \alpha_2$ and $R_S(p \to q; 0)$ is equivalent to the global restriction estimate by letting $R \to \infty$ and applying a limiting argument. Observe also that the statement for exponents $\alpha \geq n/p'$ is trivial because of Hölder's inequality, namely

$$|\hat{f}(\xi)| \leq \|f\|_1 \leq A_p R^{n/p'} \|f\|_{L^p(B(x_0,R))} .$$

Thus, the aim is to lower the value of α from the trivial value n/p' toward the ultimate aim $\alpha = 0$ for p and q belonging to the conjectured range of the restriction conjecture.

By duality, local restriction estimates are equivalent to local extension estimates, more precisely $R_S(p \to q; \alpha) \Leftrightarrow R_S^*(q' \to p'; \alpha)$ where $R_S^*(q' \to p'; \alpha)$ denotes the statement that the estimate

$$\|(F d\sigma)^\vee\|_{L^{p'}(B(x_0,R))} \leq A_{p,q,S,\alpha} R^\alpha \|F\|_{L^{q'}(S,d\sigma)} \tag{6.2}$$

holds for all smooth functions F on S, all $R \geq 1$, and all balls $B(x_0, R)$.

In the following we will focus on proving localized extension estimates, taking advantage of many phenomena not arising in the global setting. First, we observe that localizing to scale R in the spatial variable leads to a localization in frequency space on the scale R^{-1} by the uncertainty principle. More precisely, we expect F to be "blurred out" on this scale which should allow us to safely fatten up the set S to $\mathcal{N}_{R^{-1}}(S)$, the R^{-1} neighborhood of S. This is going to be made precise in the following

Lemma 6.1. *The localized extension estimate $R_S^*(q' \to p'; \alpha)$ follows from*

$$\|\check{G}\|_{L^{p'}(B(x_0,R))} \leq A_{p,q,S,\alpha} R^{\alpha-1/q} \|G\|_{L^{q'}(\mathcal{N}_{1/R}(S))} \tag{6.3}$$

whenever G is a smooth function with $\operatorname{supp} G \subseteq \mathcal{N}_{1/R}(S)$.

Remark 6.2. In the following we will make use of the following two facts.

1. For every $f \in L^1 + L^2$ with $\operatorname{supp} \hat{f} \subseteq B_0(R)$ there is a $\varphi \in \mathcal{S}(\mathbb{R}^d)$ such that $f = \varphi_R * f$ where $\varphi_R = R^d \varphi(Rx)$ (cf. Lemma D.2).
2. There are functions $0 \leq \varphi \in C_c^\infty(\mathbb{R}^d)$ with $\hat{\varphi} > 0$. To see this, take, e.g., $\psi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} \psi \subseteq B_0(C)$; then $\psi * \psi$ is supported in $B_0(2C)$ and $\mathcal{F}[\psi * \psi] = |\hat{\psi}|^2$. Thus, $\varphi = \psi * \psi$ does the job. (See also Lemma D.3.)

Proof of Lemma 6.1. Fix $R \geq 1$ and $\psi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} \psi \subseteq B(0,1)$ and $|\check{\psi}(x)| \gtrsim 1$ for all $x \in B(x_0, 1)$. Let further $G := \psi_{R^{-1}} * F d\sigma$ where $\psi_{R^{-1}}(\xi) = R^d \psi(R\xi)$. Note that this definition implies that $\operatorname{supp} G \subseteq \mathcal{N}_{R^{-1}}(S)$. Therefore, we may apply (6.3) to deduce

$$\begin{aligned} \|(F d\sigma)^\vee\|_{L^{p'}(B(x_0,R))} &\lesssim \|(F d\sigma)^\vee \check{\psi}_{R^{-1}}\|_{L^{p'}(B(x_0,R))} = \|\check{G}\|_{L^{p'}(B(x_0,R))} \\ &\leq A_{p,q,S,\alpha} R^{\alpha-1/q} \|G\|_{L^{q'}(\mathcal{N}_{R^{-1}}(S))}\,. \end{aligned}$$

Thus, it suffices to show

$$\|\psi_{R^{-1}} * (F d\sigma)\|_{L^{q'}(\mathbb{R}^d)} \lesssim R^{1/q} \|F\|_{L^{q'}(S, d\sigma)}\,.$$

For $q' = 1$, the above estimate follows immediately from Young's inequality, so by interpolation it suffices to prove the estimate for $q' = \infty$, i.e., we are left to show

$$\|\psi_{R^{-1}} * (F d\sigma)\|_\infty \lesssim R \|F\|_{L^\infty(S)}\,.$$

By Hölder's inequality, it suffices to show

$$\int_S |\psi_{R^{-1}}(\xi - \eta)| d\sigma(\eta) \lesssim R \tag{6.4}$$

uniformly in $\xi \in \mathbb{R}^d$. Heuristically, it is clear why (6.4) is true because the support of the integrand intersects S on at most a R^{d-1} cap but $\psi_{R^{-1}}$ is an L^1-scaling invariant function, i.e., the integral should be of order R. To make this argument rigorous, we will in fact prove the more general statement that whenever $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $S \subseteq \mathbb{R}^d$ is any compact hypersurface (no curvature assumption is needed!), one has

$$I(\xi) := R^n \int_S \frac{d\sigma(\eta)}{(1 + R|\xi - \eta|)^d} \lesssim R$$

uniformly in ξ for $R \gg 1$ (where we used the rapid decay of the integrand). Decomposing $S = \bigcup_{k=-1}^\infty S_k(\xi)$ with $S_k(\xi) = A_k(\xi) \cap S$ where

$$A_{-1}(\xi) := \{\eta \in \mathbb{R}^d : R|\xi - \eta| \leq 1\} \quad \text{and} \quad A_k(\xi) := \{\eta \in \mathbb{R}^d : 2^k \leq R|\xi - \eta| \leq 2^{k+1}\}\,,$$

one rewrites

$$I(\xi) = R^d \sum_{k=-1}^\infty \int_{S_k(\xi)} \frac{d\sigma(\eta)}{(1 + R|\xi - \eta|)^d}\,.$$

Now, due to the dimensionality of S, one has for any $r > 0$,

$$\sigma(B(\xi, r) \cap S) \lesssim r^{d-1}\,.$$

Indeed, this estimate is obvious for large r, whereas for $0 < r < 1$, the surface is essentially flat, i.e., $B(\xi, r) \cap S$ resembles a disk of radius r, thereby also leading to the above estimate. Thus,

$$\sigma(S_{-1}) \lesssim R^{-(d-1)} \quad \text{and} \quad \sigma(S_k) \lesssim (2^k R^{-1})^{d-1}$$

and so

$$I(\xi) \lesssim R^d \sum_{k=-1}^{\infty} (2^k R^{-1})^{d-1} \cdot 2^{-kd} \leq R\,,$$

thereby concluding the proof. □

In fact, the estimates (6.2) and (6.3) are equivalent, although the converse implication will not be used in the present section but will be referred to later.

Lemma 6.3. *The local extension estimate* $R_S^*(q' \to p'; \alpha)$ *implies* (6.3) *for all smooth functions* G *supported in* $\mathcal{N}_{1/R}(S)$.

Proof. Without loss of generality (by the translation and rotation invariance of the problem togehter with the triangle inequality), we may assume supp $G \subseteq \mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1}) \cap B_0(1/2)$. In particular, supp G is contained in the disjoint union of vertical translates $\mathbb{P}^{d-1}_\zeta := \mathbb{P}^{d-1} + (0, \zeta)$ of the paraboloid where ζ ranges over $(-R^{-1}, R^{-1}) \subset \mathbb{R}$. By Fubini's theorem and a change of variables, we have

$$\check{G}(x) = \int_{|\zeta| \leq R^{-1}} d\zeta \int_{\xi' \in [-1,1]^{d-1}} d\xi' \; G(\xi', \xi'^2) \mathrm{e}^{2\pi i x \cdot (\xi', \xi'^2)} = \int_{|\zeta| \leq R^{-1}} d\zeta \; (G|_{\mathbb{P}^{d-1}_\zeta} d\sigma_\zeta)^\vee(x)\,,$$

where $d\sigma_\zeta$ denotes the euclidean surface measure on $\mathbb{P}^{d-1}_\zeta$.

Now, assuming that the local extension estimate $R_S^*(q' \to p'; \alpha)$ holds, then it follows from translational invariance that

$$\|(G d\sigma_\zeta)^\vee\|_{L^{p'}(B_0(R))} \lesssim_\alpha R^\alpha \|G\|_{L^{q'}(\mathbb{P}^{d-1}_\zeta)} \quad \text{for all } \zeta\,.$$

Combining this estimate with Minkowski's inequality, we infer

$$\|\check{G}\|_{L^{p'}(B_0(R))} \leq \int_{|\zeta| \leq R^{-1}} d\zeta \; \|(G d\sigma_\zeta)^\vee\|_{L^{p'}(B_0(R))} \lesssim_\alpha R^\alpha \int_{|\zeta| \leq R^{-1}} d\zeta \; \|G|_{\mathbb{P}^{d-1}_\zeta}\|_{L^{q'}(\mathbb{P}^{d-1}_\zeta)}$$

and Hölder's inequality bounds the latter by

$$R^{\alpha - 1/q} \left(\int_{|\zeta| \leq R^{-1}} d\zeta \; \|G|_{\mathbb{P}^{d-1}_\zeta}\|^{q'}_{L^{q'}(\mathbb{P}^{d-1}_\zeta)} \right)^{1/q'} = R^{\alpha - 1/q} \|G\|_{L^{q'}(\mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1}))}$$

which concludes the proof. □

Obviously, the corresponding statements also hold for the restriction problem by duality, i.e.,

$$\|\hat{f}\|_{L^q(\mathcal{N}_{1/R}(S))} \leq A_{p,q,S,\alpha} R^{\alpha - 1/q} \|f\|_{L^p(B(x_0,R))} \tag{6.5}$$

for all test functions f on $B(x_0, R)$. In fact, this formulation reveals the restriction estimate $R_S(2 \to 2; 1/2)$ for smooth compact hypersurfaces S by Plancherel's theorem. (This estimate can also be obtained from the Agmon–Hörmander theorem or from the frequency localized Sobolev trace lemma.)

The obvious question now is of course how to convert local restriction estimates to global estimates. The key tool to do so is exploiting the decay of the Fourier transform $(d\sigma)^\vee$. Indeed, suppose we have a decay estimate of the form

$$|(d\sigma)^\vee(x)| \lesssim (1 + |x|)^{-\rho}$$

for some $\rho > 0$. (For hypersurfaces with everywhere non-vanishing Gaussian curvature, $\rho = (d-1)/2$, see e.g. [167, Chapter VIII, §3.1, Theorem 1]). Then, the contributions to the global restriction estimate (2.1) coming from *widely separated portions of physical space* will be almost orthogonal (in Fourier space). To make this intuition precise, suppose $R \geq 1$ and $B(x_0, R)$ and $B(x_1, R)$ are two balls which are separated by at least a distance of R. If f_j is supported on $B(x_j, R)$ $(j = 0, 1)$, then $\hat{f}_0|_S$ and $\hat{f}_1|_S$ will be almost orthogonal, namely

$$\begin{aligned} | < \hat{f}_0|_S, \hat{f}_1|_S >_{L^2(S,d\sigma)} | &= | < \hat{f}_0 d\sigma, \hat{f}_1 >_{L^2(\mathbb{R}^d)} | = | < f_0 * (d\sigma)^\vee, f_1 >_{L^2(\mathbb{R}^d)} | \\ &\lesssim R^{-\rho} \|f_0\|_{L^1(B(x_0,R))} \|f_1\|_{L^1(B(x_1,R))} \end{aligned} \tag{6.6}$$

where we used the decay assumption on $(d\sigma)^\vee(x)$ appearing in the convolution and the fact that the supports of f_0 and f_1 are separated by at least R. Put differently, the almost orthogonality in Fourier space means that distant balls in physical space do not interact much with each other.

The Tomas–Stein argument (for $R_S(2(\rho+1)/(\rho+2) \to 2)$) uses orthogonality on $L^2(S, d\sigma)$, and at first glance it seems that it can only applied to obtain restriction theorems $R_S(p \to q)$ when $q = 2$. However, Bourgain [15, 27] observed that the same type of orthogonality arguments, exploiting the decay of $(d\sigma)^\vee$, can also be used to obtain restriction estimates which are not L^2 based, albeit with some inefficiencies due to the use of non-L^2 orthogonality estimates.

Theorem 6.4. *Let ρ be as above. If $R_S(p \to q; \alpha)$ holds for some $\rho + 1 > \alpha q$, then we have $R_S(\tilde{p} \to \tilde{q})$ whenever*

$$\tilde{q} > 2 + \frac{q}{\rho + 1 - \alpha} \quad \text{and} \quad \frac{\tilde{p}}{\tilde{q}} < 1 + \frac{q}{p(\rho + 1 - \alpha q)} .$$

The ideas of the above theorem were extended by Tao [181, Theorem 1.2]. The proof will be given in Appendix A, see Theorem A.1.

Theorem 6.5. *Let ρ be as above. If $R_S(p \to p; \alpha)$ holds for some $p < 2$ and $0 < \alpha \ll 1$, then one has $R_S(q \to q)$ whenever*

$$\frac{1}{q} > \frac{1}{p} + \frac{A_\rho}{\log(1/\alpha)} .$$

Although Bourgain's theorem is more efficient for most values of α, the latter theorem is superior because it does not lose any exponents in the limit $\alpha \to 0$. In particular, we have the following consequence. If $R_S(p \to p; \varepsilon)$ holds for all $\varepsilon > 0$, then $R_S(p - \varepsilon \to p - \varepsilon)$ is also true for every $\varepsilon > 0$. (The converse statement follows easily from interpolation). Thus, one can convert a local restriction estimate with ε losses to a global estimate, where the ε loss has been transferred to the exponents. This is a prime example of an *ε-removal lemma* which is a common in this theory. These arguments show that the restriction conjecture for the paraboloid in fact states that for all $\varepsilon > 0$, the inequality

$$\|(F d\sigma)^\vee\|_{L^{2d/(d-1)}(B(x_0,R))} \leq A_{d,\varepsilon} R^\varepsilon \|F\|_{L^{2d/(d-1)}(\mathbb{P}^{d-1})}$$

holds for a suitable class of functions F on $\mathbb{P}^{d-1}$. (Note that it makes sense to consider restriction estimates at the endpoint $p' = q' = 2d/(d-1)$ in the local setting. This is another advantage of the localized setup).

7. Multilinear restriction estimates

The following ideas will be of interest of their own but also very useful to understand Bourgain and Demeter's proof of decoupling estimates. The central theme of the analysis will be the multi-linear approach, in particular the multi-linear restriction theorem of Bennett, Carbery, and Tao [7].

We start with a motivation for bilinear restriction estimates and show in particular how curvature in the linear world is translated to transversality in the multilinear world. As a striking example of the power of bilinear techniques is the complete proof by Córdoba and Fefferman of the restriction conjecture in two dimensions, see Subsection 7.4. Finally, we generalize these ideas to higher dimensions where the bilinear analysis will be replaced by a multilinear one.

The bilinear restriction estimate was first proved by Tao [176] building on earlier arguments of Wolff [199]. The work of Tao–Vargas–Vega [188] is perhaps the first systematic treatment of the bilinear phenomenon and its impact on the linear problem.

7.1. **Introduction.** The original motivation was the "L^4" or "bi-orthogonality" theory by what we mean that expressions like $\|f\|_{L^{p'}}$ can be calculated explicitly if p' is an even integer, and especially when $p'=4$. Indeed, in this case, we have, using Plancherel,

$$\|(Fd\sigma)^\vee\|_4^2 = \|(Fd\sigma)*(Fd\sigma)\|_2 \,.$$

That means that we reduced the restriction estimate $R_S^*(q'\to 4)$ (which usually crucially depends on oscillations and cancellations) to the pure size estimate

$$\|(Fd\sigma)*(Fd\sigma)\|_2 \lesssim_q \|F\|_{L^{q'}(S,d\sigma)}^2 \,,$$

which can be proven or disproven using more direct methods.

As an example, consider $d=2$ and $S=\mathbb{S}^1$, the circle. In this case, there is a logarithmic divergence in the above estimate because $d\sigma * d\sigma$ blows up like $|x|^{-1/2}$ on the circle $\{x\in\mathbb{R}^2 : |x|=2\}$ of radius 2. Localizing in physical space to a disk of radius R shows that one can easily prove the modified estimate

$$\|G*G\|_{L^2(\mathbb{R}^2)} \lesssim_q (\log R)^{1/2} R^{-3/2} \|G\|_{L^4(\mathcal{N}_{1/R}(S))}^2$$

for all $R\geq 1$ and all G with $\mathrm{supp} G \subseteq \mathcal{N}_{1/R}(S)$. Comparing this with the general localized restriction estimate (6.3) shows that this is just the restriction estimate $R_S^*(4\to 4;\varepsilon)$ for any $\varepsilon>0$. Thus, using the ε-removal lemma (Theorem 6.5), we obtain the optimal restriction estimate for the circle. Note that this was already proven by Zygmund [205] using more direct methods.

7.2. **The importance of transversality.** At first glance, this approach seems to be restricted to L^4 because of Plancherel's theorem. However, one can partially extend those ideas to other exponents p'. The main point is that the linear estimate

$$\|(Fd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)} \lesssim_{p,q,S} \|F\|_{L^{q'}(S,d\sigma)}$$

is equivalent (by squaring) to the quadratic estimate

$$\|(Fd\sigma)^\vee(Fd\sigma)^\vee\|_{L^{p'/2}(\mathbb{R}^d)} \lesssim_{p,q,S} \|F\|_{L^{q'}(S,d\sigma)}^2$$

which one can depolarize as the bilinear estimate

$$\|(F_1 d\sigma)^\vee(F_2 d\sigma)^\vee\|_{L^{p'/2}(\mathbb{R}^d)} \lesssim_{p,q,S} \|F_1\|_{L^{q'}(S,d\sigma)}\|F_2\|_{L^{q'}(S,d\sigma)} \,.$$

In such an estimate the worst case typically occurs if both F_1 and F_2 are concentrated on the same small cap on S. This is just the situation in Knapp's example (Subsection 3.2).

We saw that the basic idea is to rewrite the desired *linear* restriction estimate as a *bilinear* restriction estimate which in turn is a special case of the more general estimate

$$\|(F_1 d\sigma_2)^\vee(F_2 d\sigma_2)^\vee\|_{L^{p'/2}(\mathbb{R}^d)} \lesssim_{p,q,S_1,S_2} \|F_1\|_{L^{q'}(S_1,d\sigma_1)}\|F_2\|_{L^{q'}(S_2,d\sigma_2)} \,. \tag{7.1}$$

Here, S_1 and S_2 is a pair of smooth hypersurfaces, equipped with surface measures $d\sigma_1$ and $d\sigma_2$, respectively. Moreover, F_1 and F_2 are smooth and supported on S_1, respectively S_2. We will denote by $R_{S_1,S_2}^*(q'\times q'\to p'/2)$ the statement that (7.1) holds.

By the above discussion, $R^*_{S,S}(q' \times q' \to p'/2)$ is of course equivalent to $R^*_S(q' \to p')$. That means that bilinear restriction estimates are more general than linear ones, i.e., there are bilinear estimates that cannot be inferred from linear ones. Consider the following

Example 7.1. Let $S_1 = \{(\xi_1, 0) \in \mathbb{R}^2 : \xi_1 \in \mathbb{R}\}$ and $S_2 = \{(0, \xi_2) \in \mathbb{R}^2 : \xi_2 \in \mathbb{R}\}$, i.e., the coordinate axis in $\mathbb{R}^2$. Then $(F_1 d\sigma_1)^\vee(x, y) = \check{F}_1(x)$ and $(F_2 d\sigma_2)^\vee(x, y) = \check{F}_2(y)$ which means that there are in general no global linear restriction estimates $R^*_{S_j}(q' \to p')$ $(j = 1, 2)$, unless $p' = \infty$, since the Fourier transforms do not decay at infinity. On the other hand, Plancherel yields

$$\|(F_1 d\sigma_1)^\vee (F_2 d\sigma_2)^\vee\|^2_{L^2(\mathbb{R}^2)} = \int_{\mathbb{R}^2} |\check{F}_1(x)|^2 |\check{F}_2(y)|^2 \, dx \, dy = \|F_1\|^2_{L^2(\mathbb{R})} \|F_1\|^2_{L^2(\mathbb{R})} \,,$$

i.e., the bilinear restriction $R^*_{S_1,S_2}(2 \times 2 \to 2)$ holds, although the symmetrized estimates $R^*_{S_1,S_1}(2 \times 2 \to 2)$ and $R^*_{S_2,S_2}(2 \times 2 \to 2)$ are false.

The above example clearly indicates that *transversality* plays a major role in deriving bilinear restriction estimates (unlike in the linear situation where oscillations and cancellations were crucial). In Subsection 7.4 we will discuss bilinear estimates in $\mathbb{R}^2$ in much more detail.

Let us instead now discuss a higher-dimensional analog of the above theme. We say that two smooth hypersurfaces S_1 and S_2 are *transversal to each other*, if the set of unit normals of S_1 is separated by some non-zero distance from the set of unit normals of S_2.

Proposition 7.2. *Let S_1 and S_2 be two smooth hypersurfaces which are transversal to each other. Then, the restriction estimate $R^*_{S_1,S_2}(2 \times 2 \to 2)$ holds.*

Proof. By Plancherel, it suffices to prove the convolution estimate

$$\|(F_1 d\sigma_1) * (F_2 d\sigma_2)\|_{L^2(\mathbb{R}^d)} \lesssim_{S_1,S_2} \|F_1\|_{L^2(S_1,d\sigma)} \|F_2\|_{L^2(S_2,d\sigma)} \,.$$

By Cauchy–Schwarz,

$$\|(F_1 d\sigma_1) * (F_2 d\sigma_2)\|_{L^2(\mathbb{R}^d)} \lesssim \||F_1|^2 d\sigma_1 * |F_2|^2 d\sigma_2\|_{L^1(\mathbb{R}^d)} \|d\sigma_1 * d\sigma_2\|_{L^\infty}$$

and the second factor on the right side is bounded because of the transversality assumption. □

Generalizations of bilinear L^2 estimates arose already in works by Bourgain [18], Klainerman–Machedon [123], and many other authors in the context of nonlinear evolution equations. These estimates turned out to be especially useful for handling non-linearities which contain certain derivatives which create a "full norm".

7.3. **Necessary conditions for bilinear restriction estimates.** In this subsection, we will discuss necessary conditions for bilinear restriction estimates for the sphere and the paraboloid. Let S_1 and S_2 be two compact transverse subsets of $\mathbb{S}^{d-1}$ or $\mathbb{P}^{d-1}$. We already saw that bilinear estimates can be derived from linear ones, i.e., $R^*_S(q' \to p')$ yields by polarization the bilinear estimate $R^*_{S_1,S_2}(q' \times q' \to p'/2)$, whereas the converse statement is in general false. For instance, although the bilinear estimate $R_{S_1,S_2}(2 \times 2 \to 2)$ holds, the corresponding linear estimate $R_S(2 \to 4)$ is only true in three or higher dimensions. One reason for this is that there is no exact "transverse, bilinear" analog of the Knapp example. Indeed, the best known necessary conditions were derived by considering bilinear analogs of Knapp examples, see Foschi and Klainerman [82] and Tao–Vargas–Vega [188]. Namely, for $R^*_{S_1,S_2}(q' \times q' \to p'/2)$ to hold, one must have

$$p > \frac{2n}{n+1}; \quad \frac{n+2}{p'} + \frac{n}{q'} \leq n; \quad \frac{n+2}{p'} + \frac{n-2}{q'} \leq n - 1 \,. \tag{7.2}$$

This is somewhat less stringent than the condition

$$p > \frac{2n}{n+1}; \quad \frac{n+1}{p'} + \frac{n-1}{q'} \leq n-1$$

for the linear estimate $R_S^*(q' \to p')$. Nevertheless, the bilinear version of the restriction conjecture asserts that the conditions (7.2) are also sufficient. Apart from the case $d = 2$, this conjecture is still open. It is remarkable that it has been recently shown that the bilinear conjecture is (neglecting the endpoint) equivalent to the linear restriction conjecture for $\mathbb{S}^{d-1}$ and $\mathbb{P}^{d-1}$.

7.4. **Proof of the two-dimensional restriction conjecture.** Before we give a systematic description of multilinear restriction estimates, we present a proof of the full restriction conjecture in two dimensions involving bilinear restriction estimates. The presentation follows closely [179, Lecture 5]. The original proof goes back to Córdoba and Fefferman. Compare also to Fefferman [72, p. 33ff].

Recall that the desired estimate reads

$$\|\widehat{gd\sigma}\|_q \lesssim \|g\|_{L^p(\mathbb{S}^1)}$$

for $q > 4$ and $q \geq 3p'$ in $d = 2$. One of the fundamental reasons that the two-dimensional restriction conjecture is proved comparably easily is the involved exponent $q = 4$. One may be tempted to repeat the same argument in higher dimensions; however, it turns out that the results obtained do not improve upon Tomas–Stein and can even be worse.

As a first step, we note that it suffices to consider the quarter circle, thereby avoiding nuisances involving antipodal points. The conjecture for $\mathbb{S}^1$ then follows by the triangle inequality. Moreover, it suffices to consider the end-point $q = 3p'$ as the conjecture follows for higher q by interpolation involving Hölder's inequality.

By the enhanced Marcinkiewicz interpolation theorem (see, e.g., Tao [179, Lecture 2, Lemma 2.3] or Grafakos [96, Theorem 1.4.19] and Tao [186, Lecture 1, Lemma 8.5]), it would suffice to prove the restricted weak-type estimate (recall (C.1))

$$\|\widehat{\mathbf{1}_\Omega d\sigma}\|_{L^{q,\infty}} = \sup_{\lambda \geq 0} \lambda |\{|\widehat{\mathbf{1}_\Omega d\sigma}| \geq \lambda\}|^{1/q} \lesssim |\Omega|^{1/p}$$

where Ω is an arbitrary subset of the circle $\mathbb{S}^1$. Actually, we don't have to go quite this weak and will prove instead

$$\|\widehat{\mathbf{1}_\Omega d\sigma}\|_q \lesssim |\Omega|^{1/p} .$$

Now, the fundamental idea in the proof of the two-dimensional restriction estimate is to square it, i.e.,

$$\|\widehat{\mathbf{1}_\Omega d\sigma}\,\widehat{\mathbf{1}_\Omega d\sigma}\|_{q/2} \lesssim |\Omega|^{2/p} \tag{7.3}$$

and invoke Plancherel's theorem. Since $q > 4$, we have $2 < q/2 < \infty$, i.e., we are suddenly interested in estimating bilinear quantities such as

$$\|\widehat{fd\sigma}\,\widehat{gd\sigma}\|_2$$

and

$$\|\widehat{fd\sigma}\,\widehat{gd\sigma}\|_\infty$$

where f and g are some functions on $\mathbb{S}^1$. The latter quantity is easy to estimate, thanks to the trivial estimate $\|\widehat{fd\sigma}\|_\infty \lesssim \|f\|_1$. Thus,

$$\|\widehat{fd\sigma}\,\widehat{gd\sigma}\|_\infty \lesssim \|f\|_1 \|g\|_1 \tag{7.4}$$

and we are left with L^2 estimates. In general, it is hard to obtain good estimates for general f and g. However, if f and g are supported on disjoint arcs, i.e., they are somewhat *transversal* to each other, one obtains significant cancellations. This is summarized in the following

Lemma 7.3. *Suppose f and g are supported in distinct θ-arcs of $\mathbb{S}^1$, whose separation is also comparable θ. Then*

$$\|\widehat{fd\sigma}\,\widehat{gd\sigma}\|_2 \lesssim \theta^{-1/2}\|f\|_2\|g\|_2\,. \tag{7.5}$$

We give two proofs of this fact below due to Tao and Hickman–Vituri. Another exposition of Tao's proof is contained in the lecture notes of Zhang [204, Lecture 6].

Remark 7.4. One can make the definition of θ-separation more precise, especially for more general compact hypersurfaces. Namely, suppose $(S_j)_{j=1,\dots,n}$ is a family of compact hypersurfaces and denote by $\nu_j : S_j \to \mathbb{S}^{d-1}$ the associated Gauss map [9]. Then the S_j are said to be *θ-separated*, if

$$|\det(\nu_1(x_1),...,\nu_n(x_n))| \geq \theta \quad \text{whenever } x_j \in S_j \text{ for } j=1,...,n\,. \tag{7.6}$$

We will give two proofs of Lemma 7.3. The first one follows Tao's notes [179, Lecture 5, Lemma 1.2] and the second one the notes of Hickman and Vitturi [107, Lecture 3, Lemma 2].

Proof of Lemma 7.3 following Tao. By Plancherel, the assertion is equivalent to

$$\|(fd\sigma)*(gd\sigma)\|_2 \lesssim \theta^{-1/2}\|f\|_2\|g\|_2\,.$$

We verify this estimate by bilinear interpolation between

$$\|(fd\sigma)*(gd\sigma)\|_1 \lesssim \|f\|_1\|g\|_1$$

and

$$\|(fd\sigma)*(gd\sigma)\|_\infty \lesssim \theta^{-1}\|f\|_\infty\|g\|_\infty\,.$$

The first estimate is clear by Young's inequality (or Fubini's theorem).

To prove the second estimate, we assume that f and g are supported on θ-arcs I and J. We denote by $d\sigma_I$ and $d\sigma_J$ the restrictions of the surface measure to these arcs. By the pointwise estimates

$$fd\sigma \leq \|f\|_\infty d\sigma_I \quad \text{and} \quad gd\sigma \leq \|g\|_\infty d\sigma_J$$

it suffices to prove

$$\|d\sigma_I * d\sigma_J\|_\infty \lesssim \theta^{-1}$$

where

$$d\sigma_1 * d\sigma_2(A) := \int_{\mathbb{S}^1}\int_{\mathbb{S}^1} \mathbf{1}_A(\eta_1+\eta_2)\,d\sigma_1(\eta_1)\,d\sigma_2(\eta_2)$$

for any Borel set A lying in $\mathbb{S}^1$, and, if $d\sigma_1 * d\sigma_2$ is absolutely continuous with respect to Lebesgue, then (cf. [81, (1.8)])

$$d\sigma_1 * d\sigma_2(\xi) := \int_{\mathbb{S}^1}\int_{\mathbb{S}^1} \delta(\xi-\eta_1-\eta_2)\,d\sigma_1(\eta_1)\,d\sigma_2(\eta_2)\,.$$

To do so, we approximate $d\sigma_I$ by $(2\varepsilon)^{-1}\mathbf{1}_{I_\varepsilon}$ where $\varepsilon > 0$ is a small number and I_ε is the ε neighborhood of I, i.e.,

$$I_\varepsilon := \{r(\cos\theta,\sin\theta) :\ \theta \in I\,, 1-\varepsilon \leq r \leq 1+\varepsilon\}\,.$$

By the definition of induced Lebesgue measure, $d\sigma_I$ is the weak limit of such measures. Thus, it suffices to prove

$$\|\frac{1}{2\varepsilon}\mathbf{1}_{I_\varepsilon} * d\sigma_J\|_\infty \lesssim \theta^{-1} \tag{7.7}$$

[9]That is, ν_j continuously maps a point $x_j \in S_j$ to a choice of unit normal vector $\nu_j(x_j)$ to S_j at x_j.

for all sufficiently small θ, uniformly in ε. Clearly, the integral

$$\int_J \frac{1}{2\varepsilon}\mathbf{1}_{I_\varepsilon}(\xi-\eta)d\sigma_J(\eta) = \frac{1}{2\varepsilon}|\{\eta\in J:\,\xi-\eta\in I_\varepsilon\}| = \frac{1}{2\varepsilon}|(\xi+J)\cap I_\varepsilon|$$

only contributes whenever $\xi \in \eta + I_\varepsilon$ and $\eta \in J$. Thus, the convolution is supported on the set-theoretic sum of the arc J and the thickened arc I_ε. But since any translate of J intersects I_ε in an arc of length at most $\varepsilon\theta^{-1}$, the assertion follows. □

Remark 7.5. To avoid the convolution between measures in (7.7), one could also fatten $d\sigma_J$ there and show instead

$$\|\mathbf{1}_{I_\varepsilon} * \mathbf{1}_{J_\varepsilon}\|_{L^\infty(\mathbb{R}^2)} \lesssim \varepsilon^2|I|^{-1}\,,$$

which may be easier to verify. (Recall that the separation of I and J was supposed to be comparable, i.e., $\theta\sim|I|\sim|J|$). See also Remark 15.18 (especially Formula (15.22)) later for a similar computation, where the fattened arcs are, however, simple rectangles.

Related to convolution of measures is the following special instance, see Demeter [63, Lemma 1.20].

Lemma 7.6. *Let $d\sigma$ be the surface measure on $\mathbb{S}^{d-1}$. Then for each $d\geq 2$ the measure $d\sigma * d\sigma$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$, i.e., $d\sigma * d\sigma = F\,d\xi$ for some integrable F. Moreover* $\mathrm{supp}F\subseteq B_0(2)$ *and satisfies for a.e. ξ*

$$|F(\xi)| \lesssim \begin{cases} |\xi|^{-1}\,, & 0<|\xi|<1 \\ (2-|\xi|)^{(d-3)/2}\,, & 1\leq|\xi|\leq 2\,. \end{cases}$$

Recall that, like convolution of functions, convolutions of measures are supported on the Minkowski sum of their supports, i.e.,

$$\mathrm{supp}(\sigma_{S^{d-1}} * \sigma_{S^{d-1}}) \subseteq \mathrm{supp}(\sigma_{S^{d-1}}) + \mathrm{supp}(\sigma_{S^{d-1}}) = \{x+y:\, x,y\in\mathbb{S}^{d-1}\} \subseteq \mathbb{R}^d\,.$$

For an explicit formula, see also the survey by Foschi and Oliveira e Silva [81, (3.2)], namely

$$(\sigma_{S^{d-1}} * \sigma_{S^{d-1}})(\xi) = \frac{|\mathbb{S}^{d-2}|}{|\xi|}\left(1-\frac{|\xi|^2}{4}\right)_+^{\frac{d-3}{2}}\,. \tag{7.8}$$

This shows that the $|\xi|^{-1}$ singularity in the lemma is in fact necessary, and hence the Radon–Nikodym derivative $d\sigma * d\sigma$ with respect to Lebesgue exists, but is not bounded. Essentially this is due to the large symmetry of $\mathbb{S}^{d-1}$ which leads to the fact that the origin can be represented in multiple ways by $\xi+\eta$ where $\xi,\eta\in\mathbb{S}^{d-1}$. Heuristically, this is another reason why we split $\mathbb{S}^1$ into multiple chunks so that "most of the time different arcs cannot too badly with each other".

Proof. Let S_ε^{d-1} be the ε-neighborhood of S^{d-1} and let $\sigma_\varepsilon := \varepsilon^{-1}\mathbf{1}_{S_\varepsilon^{d-1}}$. Then $\sigma_\varepsilon\,d\xi \rightharpoonup d\sigma$ as $\varepsilon\to 0$. Note that

$$\sigma_\varepsilon * \sigma_\varepsilon(\xi) = \varepsilon^{-2}\int_{\mathbb{R}^d}\mathbf{1}_{S_\varepsilon^{d-1}}(\xi-\eta)\mathbf{1}_{S_\varepsilon^{d-1}}(\eta)\,d\eta = \varepsilon^{-2}|S_\varepsilon^{d-1}\cap(\xi+S_\varepsilon^{d-1})|\,.$$

The right side is zero for $|\xi|>2$. Since $S_\varepsilon^{d-1}\cap(\xi+S_\varepsilon^{d-1})$ is a body of revolution, its volume is at most a constant multiple of the area of the cross section $S_\varepsilon^1\cap((r,0)+S_\varepsilon^1)$ with $r=|\xi|$.

Now suppose $r\leq 1$. Then note that any $y=(y_1,y_2)\in S_\varepsilon^1\cap((r,0)+S_\varepsilon^1)$ satisfies

$$\begin{aligned} 1-2\varepsilon \leq y_1^2+y_2^2 \leq 1+3\varepsilon\,, &\quad \text{since } y\in S_\varepsilon^1 \text{ and } y\in((r,0)+S_\varepsilon^1) \\ 1-2\varepsilon \leq (y_1-r)^2+y_2^2 \leq 1+3\varepsilon\,, &\quad \text{since } (y_1-r,y_2)\in S_\varepsilon^1 \end{aligned}$$

and thus also (combining the first lower bound $y_1^2+y_2^2 \geq 1-2\varepsilon$ with the second upper bound $-2y_1 r + r^2 + y_1^2 + y_2^2 \leq 1+3\varepsilon$),

$$|2y_1 - r| \leq \frac{5\varepsilon}{r}\,.$$

This means that the horizontal projection of $S_\varepsilon^1 \cap ((r,0)+S_\varepsilon^1)$ sits inside an interval of length $5\varepsilon/r$. Since $r \leq 1$ the vertical slices of $S_\varepsilon^1 \cap ((r,0)+S_\varepsilon^1)$ have length $\lesssim \varepsilon$. Using Fubini, we find that $|S_\varepsilon^1 \cap ((r,0)+S_\varepsilon^1| \lesssim \varepsilon^2/r$. Thus, if $|\xi| \leq 1$, then

$$\sup_{\varepsilon\in(0,1)} \sigma_\varepsilon * \sigma_\varepsilon(\xi) \lesssim |\xi|^{-1}\,.$$

Finally a similar computation shows that if $1 \leq |\xi| \leq 2$, then

$$\sup_{\varepsilon\in(0,1)} \sigma_\varepsilon(\xi) * \sigma_\varepsilon(\xi) \lesssim (2-|\xi|)^{\frac{d-3}{2}}\,.$$

Since $\sigma_\varepsilon * \sigma_\varepsilon \, d\xi \rightharpoonup d\sigma * d\sigma$, the proof is concluded. □

Proof of Lemma 7.3 following Hickman–Vitturi. For $\xi \in \mathbb{S}^1$ we can approximate the circle locally by a parabola which can be parameterized by (t,t^2) for $t \in \mathbb{R}$. Now, since we are assuming that the two arcs of length θ are only θ-separated and θ is supposed to be very tiny, we can assume that these arcs are actually θ-transverse caps on the one-dimensional parabola $\mathbb{P}^1$. So, let $I_1, I_2 \subseteq [0,1]$ be the two intervals parameterizing these caps. By the transversality condition, I_1 and I_2 are $\mathcal{O}(\theta)$-separated. Denoting $g_1 = f$ and $g_2 = g$, we observe

$$\begin{aligned}\prod_{j=1}^2 \widehat{g_j d\sigma}(x) &= \int_{I_2}\int_{I_1}\prod_{j=1}^2 g_j(t_j,t_j^2) \mathrm{e}^{2\pi i[x_1(t_1+t_2)+x_2(t_1^2+t_2^2)]}\,dt_1\,dt_2\\ &= 2^{-1}\iint_D \prod_{j=1}^2 g_j(t_j(u),t_j(u)^2)|t_1(u)-t_2(u)|^{-1}\mathrm{e}^{2\pi i x\cdot u}\end{aligned}$$

where we have applied the change of variables $u_1 = t_1+t_2$ and $u_2 = t_1^2+t_2^2$. [10] The latter is the Fourier transform of a bivariate function and so, by Plancherel, we have

$$\begin{aligned}\|\prod_{j=1}^2 \widehat{g_j d\sigma}\|_2^2 &= 2^{-2}\iint_D \prod_{j=1}^2 |g_j(t_j(u),t_j(u)^2)|^2 |t_1(u)-t_2(u)|^{-2}\,du\\ &= 2^{-1}\int_{I_1}\int_{I_2} |g_j(t_j,t_j^2)|^2|t_1-t_2|^{-1}\,dt_1\,dt_2\end{aligned}$$

The result now follows from $|t_1-t_2| \gtrsim \theta$ which is a consequence of the separation hypothesis. □

Remark 7.7. Note that this argument can be generalized to prove n-linear restriction estimates for θ-separated pieces of the *moment curve* $t \mapsto (t,t^2,\ldots,t^n)$. Here, the Jacobian arising from the above indicated change of variables is a (scalar multiple of a) Vandermonde determinant and one can use the same argument as of the footnote to prove the injectivity of the mapping, now invoking the Newton–Girard formulae.

---

[10] To see that this change of variables is valid on $I_1 \times I_2$, note that if $s_j, t_j \in I_j$ (for $j=1,2$) satisfy

$$s_1+s_2 = t_1+t_2 \quad\text{and}\quad s_1^2+s_2^2 = t_1^2+t_2^2\,.$$

Then it follows from the formula $2ab = (a+b)^2 - (a^2+b^2)$ that $s_1 s_2 = t_1 t_2$. Consequently, by comparing coefficients, we see that $\prod_{j=1}^2 (z-t_j)$ and $\prod_{j=1}^2(z-s_j)$ define the same polynomial (here, z is a single complex variable) and hence the t_j equal the s_j, up to permutation. The separation of the intervals now implies $t_j = s_j$ for $j=1,2$.

To prove (7.3), we have to piece all this together. However, we cannot simply interpolate between (7.4) and (7.5) because of the support restrictions in Lemma 7.3. Therefore, we will split the left side of (7.3) into pieces in order to exploit (7.5).

To this end, we will use the *Whitney decomposition.* For every $n \geq 0$, we divide $\mathbb{S}^1$ into 2^n equal arcs, so that each arc at stage n has exactly two children at stage $n+1$. We denote the set of all arcs at stage n by A_n. We say that two arcs $I, J \in A_n$ from the same stage $n > 1$ are *close*, if they are not adjacent, but their parents are adjacent. In this case, we write $I \sim J$. Note that for each I there are only $\mathcal{O}(1)$ arcs J which are close to J.

Remark 7.8. Here we see that the non-vanishing curvature condition is crucial in the linear problem as it allows us to find sufficiently many transverse pairs of arcs in the bilinear problem.

For every $x \neq y$ on $\mathbb{S}^1$, there is exactly one pairs of arcs I, J containing x and y respectively such that $I \sim J$. This implies (by imagining the following formula in Fourier space)

$$\widehat{\mathbf{1}_\Omega d\sigma}\,\widehat{\mathbf{1}_\Omega d\sigma} = \sum_{I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J} = \sum_{n>1}\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\,.$$

Remark 7.9. This decomposition is somewhat special to the bilinear perspective and so far, there seems to be no known satisfactory way to duplicate this in a linear setting.

We are now ready to plug this decomposition into (7.3). To deal with the n summation, we simply use the triangle inequality to obtain

$$\|\widehat{\mathbf{1}_\Omega d\sigma}\,\widehat{\mathbf{1}_\Omega d\sigma}\|_{q/2} \lesssim \sum_{n>1} \|\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\|_{q/2}\,.$$

We will estimate the $L^{q/2}$ norm by interpolating between estimates on the L^∞ and the L^2 norm and we begin with the former. By the triangle inequality and (7.4), we obtain

$$\|\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\|_\infty \lesssim \sum_{I,J\in A_n : I\sim J} |\Omega\cap I||\Omega\cap J|\,.$$

Although there are no Fourier transforms appearing on the right side anymore, a more tractable dependence on Ω or factors of 2^{-n} would be desirable. Fortunately, similar crude estimates will do the trick. Clearly, we may estimate $|\Omega \cap J| \lesssim 2^{-n}$ at stage n. But since there are only $\mathcal{O}(1)$ arcs J for each I, we obtain on the one hand

$$\|\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\|_\infty \lesssim \sum_{I\in A_n} |\Omega\cap I|\cdot 2^{-n} = 2^{-n}|\Omega|\,.$$

Alternatively, we may simply lift the restriction $I \sim J$ on the summation and obtain

$$\|\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\|_\infty \lesssim \left(\sum_{I\in A_n} |\Omega\cap I|\right)\left(\sum_{J\in A_n} |\Omega\cap J|\right) = |\Omega|^2\,.$$

Combining the last two estimates therefore shows

$$\|\sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}\|_\infty \lesssim |\Omega|\min\{|\Omega|, 2^{-n}\}\,. \tag{7.9}$$

Thus, we are left with the L^2 estimate. This time the triangle inequality is a bad idea as there are lots of oscillations and orthogonality present that should be exploited more effectively. The following observation of Fefferman is fundamental for what comes next.

As $I \sim J$ vary, the functions $\widehat{\mathbf{1}_\Omega d\sigma_I}\,\widehat{\mathbf{1}_\Omega d\sigma_J}$ have Fourier transform supports which are essentially disjoint which means that the functions themselves are essentially orthogonal. This is just done by computing the set theoretic sums of I and J and computing. Because of this (almost) orthogonality [11], we thus have

$$
\begin{aligned}
&\| \sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\, \widehat{\mathbf{1}_\Omega d\sigma_J} \|_2 \\
&\lesssim \left(\sum_{I,J\in A_n : I\sim J} \| \widehat{\mathbf{1}_\Omega d\sigma_I}\, \widehat{\mathbf{1}_\Omega d\sigma_J} \|_2^2 \right)^{1/2} \\
&\lesssim 2^{n/2} \left(\sum_{I,J\in A_n : I\sim J} |\Omega\cap I||\Omega \cap J| \right)^{1/2},
\end{aligned}
$$

where we used Lemma 7.3 with $\theta = 2^{-n}$ in the final inequality. By the same arguments as before, we estimate the sum over the close arcs and obtain

$$
\| \sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\, \widehat{\mathbf{1}_\Omega d\sigma_J} \|_2 \lesssim 2^{n/2} (|\Omega| \, \min\{|\Omega|, 2^{-n}\})^{1/2} .
$$

Combining this with (7.9) by Hölder's inequality, we thus have

$$
\| \sum_{I,J\in A_n : I\sim J} \widehat{\mathbf{1}_\Omega d\sigma_I}\, \widehat{\mathbf{1}_\Omega d\sigma_J} \|_{q/2} \lesssim 2^{2n/q} (|\Omega| \, \min\{|\Omega|, 2^{-n}\})^{1-2/q} .
$$

Finally, summing over n, we obtain

$$
\| \widehat{\mathbf{1}_\Omega d\sigma}\, \widehat{\mathbf{1}_\Omega d\sigma} \|_{q/2} \lesssim \sum_{n>1} 2^{2n/q} (|\Omega| \, \min\{|\Omega|, 2^{-n}\})^{1-2/q} ,
$$

where the right side can be computed (by considering $2^{-n} > |\Omega|$ and $2^{-n} > |\Omega|$ separately) to be $|\Omega|^{1-2/q} = |\Omega|^{2/p}$, which was desired.

Remark 7.10. A quite similar argument can be used to prove the Bochner–Riesz conjecture in $d = 2$.

One of the key innovations here was the bilinear approach. Unfortunately, one cannot apply the above argument directly to higher dimensions unless $q/2 \geq 2$. (As one can check, these cases are already taken care of by the Tomas–Stein estimate.) Nevertheless, the bilinear approach was quite useful in higher dimensions, and in fact all the best results on the restriction conjecture and related problems has come from precisely such an approach.

Remark 7.11. The original proof of Córdoba and Fefferman did not take such an explicitly bilinear approach, and was more elegant; however, it was less obvious whether any of the ideas could be extended to other dimensions and exponents.

7.5. **From bilinear to linear.** The most valuable feature of the bilinear restriction conjecture is the fact that it implies the linear restriction conjecture. For technical reasons, consider only compact subsets of the paraboloid.

Proposition 7.12 ([188]). *Let $S \subseteq \mathbb{P}^{d-1}$ be compact and S_1 and S_2 transversal subsets of S. If $q > 2d/(d-1)$ and $q \geq p'(d+1)/(d-1)$, and the conjectured bilinear inequality*

$$
\| \widehat{f_1\, d\sigma} \widehat{f_2\, d\sigma} \|_{L^{\tilde q/2}(\mathbb{R}^d)} \lesssim \|f_1\|_{L^{\tilde p}(S_1)} \|f_2\|_{L^{\tilde p}(S_2)}
$$

[11] One can obtain perfect orthogonality by only considering, say, every tenth pair (I,J) and then add up the ten smaller sums by the triangle inequality.

holds for all $(\tilde{p},\tilde{q})$ in a neighborhood of (p,q), then the conjectured linear inequality

$$\|\widehat{f\,d\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|f\|_{L^p(d\sigma)}$$

holds.

We first follow Bennett [6, p. 7-8]. See Tao–Vargas–Vega [188] for the original argument (Theorem 2.2 for the global and Theorem 4.1 for the local restriction estimates) and Bourgain–Guth [31] for a simpler argument. For a textbook treatment see Demeter [63, Chapter 7]. In the second subsection we present an argument relying on parabolic rescaling which is borrowed from Demeter [63, Chapter 4].

7.5.1. *Bourgain–Guth method.* We present the argument in a such a way that it may be adapted to a more general multilinear setting.

Sketch of proof of Proposition 7.12. We show the extension estimate $\|\widehat{f\,d\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|f\|_{L^p(d\sigma)}$ in the range $p = q > 2d/(d-1)$. This special case is readily seen to imply the linear restriction conjecture on the interior of the full conjectured range of Lebesgue exponents.

Let $\{S_\alpha\}$ be a partition of S by patches of diameter approximately $1/K$ and write

$$f = \sum_\alpha f_\alpha\,, \quad f_\alpha := f\chi_{S_\alpha}\,.$$

By linearity, $\widehat{f\,d\sigma} = \sum_\alpha \widehat{f_\alpha\,d\sigma}$. The key observation is the following inequality, see Bourgain–Guth [31]. □

Proposition 7.13. *We have*

$$|\widehat{f\,d\sigma}(x)|^q \lesssim K^{2(d-1)q} \sum_{S_{\alpha_1},S_{\alpha_2}} |\widehat{f_{\alpha_1}\,d\sigma}(x)\widehat{f_{\alpha_2}\,d\sigma}(x)|^{q/2} + \sum_\alpha |\widehat{f_\alpha\,d\sigma}|^q\,, \tag{7.10}$$

where the sum in S_{α_1} and S_{α_2} is restricted to $1/K$-transversal pairs S_{α_1} and S_{α_2}, i.e., $|v_1 \wedge v_2| \geq 1/K$ for all choices of unit normal vectors v_1, v_2 to $S_{\alpha_1}, S_{\alpha_2}$, respectively.

Remark 7.14. Recall that for $v_1, ..., v_k \in \mathbb{R}^n$ with $k \leq n$ the wedge product $v_1\wedge...\wedge v_k$ belongs to the k-th exterior power of $\mathbb{R}^n$, denoted by $\Lambda^k(\mathbb{R}^n)$ with dimension $\dim(\Lambda^k(\mathbb{R}^n)) = \binom{n}{k}$. Suppose $e_1, ..., e_{\binom{n}{k}}$ is a basis of $\Lambda^k(\mathbb{R}^n)$ and that $v_1 \wedge ... \wedge v_k \equiv \sum_{\ell=1}^{\binom{n}{k}} a_\ell e_\ell$. Then

$$|v_1 \wedge ... \wedge v_k| = \left(\sum_{\ell=1}^{\binom{n}{k}} |a_\ell|^2\right)^{1/2}. \tag{7.11}$$

In the special case where $k = n$ and the components of the vector v_j are denoted by $\{v_{j,\ell}\}_{\ell=1}^n$, we have

$$|v_1 \wedge ... \wedge v_k| = \left|\det \begin{pmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,n} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n,1} & v_{n,2} & \cdots & v_{n,n} \end{pmatrix}\right| \tag{7.12}$$

Proof of Proposition 7.13. This essentially amounts to an application of the elementary abstract inequality

$$\|a\|_{\ell^1(\mathbb{Z}_N)}^q \lesssim N \sum_{j\neq k} |a_j a_k|^{q/2} + \|a\|_{\ell^q(\mathbb{Z}_N)}^q$$

for finite sequences of real numbers a. □

Continuation of sketch of proof of Proposition 7.12. Assuming the truth of Proposition 7.13 and integrating in x, we obtain `[where does the` $K^{2(d-1)q}$ `come from?]`

$$\|\widehat{f\,d\sigma}\|_q^q \lesssim K^{2(d-1)q} \sum_{S_{\alpha_1},S_{\alpha_2}} \|\widehat{f_{\alpha_1}\,d\sigma}\widehat{f_{\alpha_2}\,d\sigma}\|_{q/2}^{q/2} + \sum_\alpha \|\widehat{f_\alpha\,d\sigma}\|_q^q\,, \tag{7.13}$$

which, because of the terms $\|\widehat{f_\alpha\,d\sigma}\|_q^q$ appearing on the right side, strongly suggests the viability of a bootstrapping argument. To this end, let $\mathcal{C} = \mathcal{C}(R)$ denote the smallest constant in the inequality $\|\widehat{f\,d\sigma}\|_{L^q(B_0(R))} \leq C\|f\|_q$ over all $R \gg 1$ and $f \in L^p(d\sigma)$. The only role of the parameter R here is to ensure that $\mathcal{C}$ is a-priori finite. Our goal is to show $\mathcal{C} < \infty$, uniformly in R. Because of the Fourier cut-off on S_α (which has diameter $1/K$), the hypothesis gives `[where does the` $K^{2d/q-(d-1)}$ `come from?]`

$$\|\widehat{f_\alpha\,d\sigma}\|_q \lesssim \mathcal{C}K^{2d/q-(d-1)}\,.$$

Since $2d/q-(d-1) < 0$ and $K \gg 1$, this represents a gain! Using (7.13) along with the property $\sum_\alpha \|f_\alpha\|_q^q = \|f\|_q^q$ (by Fourier disjointness), we obtain

$$\|\widehat{f\,d\sigma}\|_q^q \leq cK^{2(d-1)q} \sum_{S_{\alpha_1},S_{\alpha_2}} \|\widehat{f_{\alpha_1}\,d\sigma}\widehat{f_{\alpha_2}\,d\sigma}\|_{q/2}^{q/2} + \mathcal{C}K^{2d/q-(d-1)}\|f\|_q^q \tag{7.14}$$

for some constant c independent of K. Taking K so large that $cK^{2d/q-(d-1)} \leq 1/2$ (say), we see that it suffices to show

$$K^{2(d-1)q} \sum_{S_{\alpha_1},S_{\alpha_2}} \|\widehat{f_{\alpha_1}\,d\sigma}\widehat{f_{\alpha_2}\,d\sigma}\|_{q/2}^{q/2} \leq A(K)\|f\|_q^q\,. \tag{7.15}$$

Believing this estimate for a moment, then by definition of $\mathcal{C}$, we have $\mathcal{C} \leq cA + \mathcal{C}/2$, from which we may deduce that $\mathcal{C} < \infty$ uniformly in R. However, (7.15) is a straightforward consequence of the conjectured bilinear inequality. □

Remark 7.15. The above argument would have been equally effective if the factor $K^{2(d-1)}$ in (7.10) were replaced by any fixed power of K. As we have seen, the key feature of (7.10) is the absence of a power of K in the second "bootstrapping" term on the right side.

7.5.2. *Parabolic rescaling and bilinear to linear reduction.* Parabolic rescaling means that the affine functions

$$\mathbb{R}^{d-1} \times \mathbb{R} \ni (\xi, \xi_d) \mapsto \left(\frac{\xi - \xi_0}{\delta}, \frac{\xi_d - 2\xi_0 \cdot \xi + |\xi_0|^2}{\delta^2}\right)\,, \quad \delta > 0,\ \xi_0 \in \mathbb{R}^{d-1}$$

map the (infinite) paraboloid into intself.

Next recall that any nonsingular affine map $T(\eta) = A\eta + v$ (for some $d \times d$ matrix A) interacts with the Fourier transform via

$$\hat{G} := \hat{F} \circ T \Rightarrow G(x) = \frac{1}{\det(A)} F((A^{-1})^t x) \mathrm{e}^{-2\pi i \langle v, (A^{-1})^t x\rangle}\,.$$

Our goal is to use change of variables to convert inequalities involving functions whose Fourier support lives on or near a small cap on $\mathbb{P}^{d-1}$ into similar inequalities involving functions whose Fourier support is then spread over neighborhoods of the whole $\mathbb{P}^{d-1}$. To make this precise, we will need to measure the constants appearing in such inequalities precisely. In the context of bilinear restriction, we make the following

Definition 7.16 (Bilinear restriction constants)**.** Let $1 \leq p, q \leq \infty$ and $0 < D \leq 1$. We denote by $BR^*(q \times q \mapsto p, D)$ the smallest constant C such that for each set of cubes $\Omega_1, \Omega_2 \subseteq [-1,1]^{d-1}$ with $\mathrm{dist}(\Omega_1, \Omega_2) \geq D$ and each $f : \Omega_1 \cup \Omega_2 \to \mathbb{C}$, we have

$$\|E_{\Omega_1} f \, E_{\Omega_2} f\|_{L^{p/2}(\mathbb{R}^d)} \leq C \|f\|_{L^q(\Omega_1)} \|f\|_{L^q(\Omega_2)} .$$

We will now parabolically rescale the known bilinear restriction estimates and afterwards combine these with a Whitney decomposition to derive new linear restriction estimates.

Proposition 7.17. *Let Ω_1, Ω_2 be two cubes in $[-1,1]^{d-1}$ with side length δ and assume that the distance D between their centers satisfies $D \geq 4\delta$. Then for each $1 \leq p, q \leq \infty$ and each $f : \Omega_1 \cup \Omega_2 \to \mathbb{C}$, we have*

$$\|E_{\Omega_1} f E_{\Omega_2} f\|_{L^{p/2}(\mathbb{R}^d)} \leq D^{\frac{2(d-1)}{q'} - \frac{2(d+1)}{p}} BR^*(q \times q \mapsto p, \frac{1}{2}) \|f\|_{L^q(\Omega_1)} \|f\|_{L^q(\Omega_2)} .$$

Note that the exponent of D is non-negative when p, q are in the linear restriction range. Thus, we have extra gain as D gets smaller.

Proof. Let ξ_0 be the midpoint of the line segment joining the centers of Ω_1 and Ω_2. Define an affine transformation on $\mathbb{R}^{d-1}$ by

$$L(\xi) \equiv L_{\xi_0, D}(\xi) := \frac{\xi - \xi_0}{D} .$$

Then a simple computation shows that

$$|E_{\Omega_i} f(x', x_d)| = D^{d-1} |E_{L(\Omega_i)} f_L(D(x' + 2x_d \xi_0), D^2 x_d)| , \quad f_L := f \circ L^{-1} .$$

Note that $L(\Omega_1)$ and $L(\Omega_2)$ are now cubes in $[-1,1]^d$ that are separated by at least $1/2$ (instead of 2δ. Changing variables on the spatial side then gives

$$\begin{aligned} \|E_{\Omega_1} f E_{\Omega_2} f\|_{L^{p/2}(\mathbb{R}^d)} &= D^{2(d-1) - 2(d+1)/p} \|E_{L(\Omega_1)} f_L \, E_{L(\Omega_2)} f_L\|_{L^{p/2}(\mathbb{R}^d)} \\ &\leq D^{2(d-1) - 2(d+1)/p} \, BR^*(q \times q \mapsto p, \frac{1}{2}) \|f_L\|_{L^q(L(\Omega_1))} \|f_L\|_{L^q(L(\Omega_2))} \\ &= D^{\frac{2(d-1)}{q'} - \frac{2(d+1)}{p}} BR^*(q \times q \mapsto p, \frac{1}{2}) \|f\|_{L^q(\Omega_1)} \|f\|_{L^q(\Omega_2)} , \end{aligned}$$

which concludes the proof. □

We recall the dyadic Whitney decomposition. A dyadic interval is an interval of the form $[\ell 2^k, (\ell+1)2^k]$ with $\ell, k \in \mathbb{Z}$. A dyadic cube is the Cartesian product of dyadic intervals of equal length. If two dyadic cubes intersect, then one must be the subset of the other.

Proposition 7.18 (Dyadic Whitney decomposition)**.** *Let $S \subseteq \mathbb{R}^m$ be a closed set. Then there is a collection $\mathcal{C}$ of closed dyadic cubes Ω with pairwise disjoint interiors such that*

$$\mathbb{R}^m \setminus S = \bigcup_{\Omega \in \mathcal{C}} \Omega$$

and whose sidelength $\ell(\Omega)$ grows with the distance to S by

$$4\ell(\Omega) \leq \mathrm{dist}(\Omega, S) \leq 50\ell(\Omega) . \tag{7.16}$$

Proof. See Demeter [63, Proposition 4.3] or Tao [186, Lecture 3, Proposition 4.6]. □

We need this in the following particular case.

Corollary 7.19. *Let $d \geq 2$, then there is a collection $\mathcal{C}$ of closed cubes $\Omega = \Omega_1 \times \Omega_2 \subseteq [-1,1]^{d-1} \times [-1,1]^{d-1}$ with pairwise disjoint interiors such that*

$$[-1,1]^{2d-2} \setminus \{(\xi,\xi) : \xi \in [-1,1]^{d-1}\} = \bigcup_{\Omega \in \mathcal{C}} \Omega$$

and

$$4\ell(\Omega) \leq \mathrm{dist}(\Omega_1, \Omega_2) \leq 100\ell(\Omega)\,. \tag{7.17}$$

Observe that the lower bound in (7.17) reflects the fact that the cubes Ω do not intersect the diagonal. However, the bounds say that their side length is still comparable to the distance between the underlying Ω_1 and Ω_2. In $d = 2$ this is illustrated in the following figure.

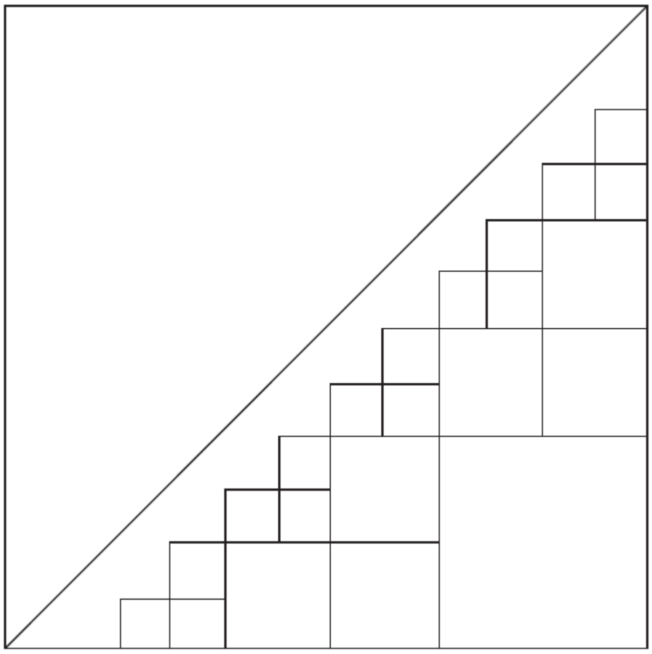

Figure 1. Dyadic Whitney decomposition of $[-1,1]^2$ in the lower triangle

Proof. It suffices to achieve a similar decomposition with $[-1,1]$ replaced with $[0,2]$ and then translate the cubes by $(-1,-1,...,-1)$. The advantage of working with $[0,2]$ is that it is already a dyadic interval.

Use the family of Proposition 7.18 with $m = 2d-2$ and $S = \{(\xi,\xi) : \xi \in \mathbb{R}^{d-1}\}$ and only keep the cubes that are inside $[0,2]^{2d-2}$. They obviously cover $[0,2]^{2d-2}$. Likewise, the bounds (7.17) follow from (7.16). $\square$

The following lemma says that when we have a sequence of functions with disjoint Fourier support, we can easily decouple their contributions to an L^s norm.

Lemma 7.20. *Let $\mathcal{R}$ be a finite collection of rectangular boxes in $\mathbb{R}^d$ with $2R \cap 2R' = \emptyset$ whenever $R \neq R' \in \mathcal{R}$. For $R \in \mathcal{R}$ let $F_R : \mathbb{R}^d \to \mathbb{C}$ be an L^s function for some $1 \leq s \leq \infty$ with $\mathrm{supp}(\hat{F_R}) \subseteq R$. Then*

$$\|\sum_R F_R\|_{L^s} \lesssim \left(\sum_R \|F_R\|_s^s\right)^{1/s}\quad, \quad 1 \leq s \leq 2$$

and

$$\|\sum_R F_R\|_{L^s} \lesssim \left(\sum_R \|F_R\|_s^{s'}\right)^{1/s'}, \quad s \geq 2$$

where the implicit constants do not depend on R.

Proof. Let $\varphi_R \in \mathcal{S}(\mathbb{R}^d)$ with $\mathbf{1}_R \leq \hat{\varphi}_R \leq \mathbf{1}_{2R}$ and $\|\varphi_R\| = 1$. Note that $F_R = F_R * \varphi_R$. Consider the operator T acting on an arbitrary family $G_{\mathcal{R}} = (G_R)_{R\in\mathcal{R}}$ of functions G_R via

$$T(G_{\mathcal{R}}) = \sum_R G_R * \varphi_R\,.$$

By orthogonality, $T : L^2(\mathbb{R}^d : \ell^2(\mathcal{R})) \to L^2(\mathbb{R}^d)$, i.e.,

$$\begin{aligned}\|T(G_{\mathcal{R}})\|_2^2 &= \int |\sum_R G_R * \varphi_R(x)|^2\,dx = \int |\sum_R \hat{G}_R\hat{\varphi}_R(\xi)|^2 = \sum_R \int |\hat{G}_R\hat{\varphi}_R(\xi)|^2 \\ &= \sum_R \int |G_R * \varphi_R(x)|^2 \leq \sum_R \|G_R\|_2^2\end{aligned}$$

and by Young's inequality, $T : L^1(\mathbb{R}^d : \ell^1(\mathcal{R})) \to L^1(\mathbb{R}^d)$, i.e.,

$$\|T(G_{\mathcal{R}})\|_1 = \int |\sum_R G_R * \varphi_R(x)|\,dx \leq \sum_R \|G_R * \varphi_R\|_1] \leq \sum_R \|G_R\|_1\,.$$

Vector-valued interpolation thus gives the first assertion. Since $T : L^\infty(\mathbb{R}^d, \ell^1(\mathcal{R})) \to L^\infty$, i.e.,

$$\|T(G_{\mathcal{R}})\|_\infty = \|\sum_R G_R * \varphi_R\|_\infty \leq \sum_R \|G_R\|_\infty\,,$$

vector-valued interpolation also gives the second assertion. □

We are now ready to assemble all the previous ingredients and prove that bilinear restriction estimates give linear ones.

Theorem 7.21. *Assume that*

$$BR^*(\infty\times\infty \mapsto p, \frac{1}{2}) < \infty$$

for some $2d(d-1) < p \leq 4$ *when* $d \geq 3$ *or for some* $p > 4$ *when* $d = 2$. *Then the linear estimate* $R^*(\infty \to p)$ *holds.*

Proof. Let $\mathcal{C}$ be a collection of closed cubes $\Omega = \Omega_1 \times \Omega_2 \subseteq [-1,1]^{d-1} \times [-1,1]^{d-1}$ as in Corollary 7.19. Let $f : [-1,1]^{d-1} \to \mathbb{C}$. Then we may write (neglecting the diagonal which has Lebesgue measure zero)

$$\begin{aligned}Ef(x)^2 &= \int_{[-1,1]^{d-1}\times[-1,1]^{d-1}} f(\xi_1)f(\xi_2)\mathrm{e}^{2\pi i x'\cdot(\xi_1+\xi_2)+2\pi i x_d(\xi_1^2+\xi_2^2)}\,d\xi_1\,d\xi_2 \\ &= \sum_{\Omega=\Omega_1\times\Omega_2\in\mathcal{C}} \int_{\Omega_1\times\Omega_2} f(\xi_1)f(\xi_2)\mathrm{e}^{2\pi i x'\cdot(\xi_1+\xi_2)+2\pi i x_d(\xi_1^2+\xi_2^2)}\,d\xi_1\,d\xi_2 \\ &= \sum_{\Omega\in\mathcal{C}} E_{\Omega_1}f(x)E_{\Omega_2}f(x)\,.\end{aligned}$$

Now for $k \geq 1$ define $\mathcal{C}_k$ to consist of these cubes in $\mathcal{C}$ whose side length is 2^{-k}. We separate these scales using the triangle inequality and obtain

$$\|Ef\|_p^2 = \|(Ef)^2\|_{p/2} \leq \sum_{k\geq 1} \|\sum_{\Omega\in\mathcal{C}_k} E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}\,. \tag{7.18}$$

Now note that as Ω ranges through $\mathcal{C}_k$, the collection of cubes $4(\Omega_1+\Omega_2)$ overlap at most C times for some C independent of k. This follows from the following two observations.

(1) The upper bound $\mathrm{dist}(\Omega_1,\Omega_2)\leq 100\ell(\Omega)=100\ell(\Omega_1\times\Omega_2)$ in (7.17) forces $\Omega_1+\Omega_2\subseteq\Omega_1+1000\Omega_1$.

(2) Each Ω_1 appears at most $\mathcal{O}(1)$ times as the first component of some $\Omega\in\mathcal{C}_k$. (This observation will allows us to exploit orthogonality within each family $\mathcal{C}_k$.)

We would now like to appply Lemma 7.20 with $s=p/2$ and $F_R=E_{\Omega_1}fE_{\Omega_2}f$. The Fourier transform of $E_{\Omega_1}fE_{\Omega_2}f$ is supported inside a rectangular box $R_\Omega\subseteq\mathbb{R}^{d-1}\times\mathbb{R}$ whose projection to $\mathbb{R}^{d-1}$ lies inside $2(\Omega_1+\Omega_2)$. But by the finite overlaps of $4(\Omega_1+\Omega_2)$ (discussed above) it follows that we can split $\mathcal{C}_k$ into $C=O(1)$ families such that the boxes $2R_\Omega$ are pairwise disjoint for Ω in each family. By applying Lemma 7.20 with $s=p/2$ to each family, we obtain

$$\|\sum_{\Omega\in\mathcal{C}_k}E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}\lesssim\left(\sum_{\Omega\in\mathcal{C}_k}\|E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}^{p/2}\right)^{2/p},\quad d\geq 3$$

respectively

$$\|\sum_{\Omega\in\mathcal{C}_k}E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}\lesssim\left(\sum_{\Omega\in\mathcal{C}_k}\|E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}^{p/(p-2)}\right)^{p-2/p},\quad d=2$$

where the implicit constants do not depend on k.

Now the lower bound in (7.17) (i.e., $\mathrm{dist}(\Omega_1,\Omega_2)\geq 4\ell(\Omega)$) allows us to apply the parabolically rescaled bilinear estimate of Proposition 7.17 to each term in the sum and obtain

$$\|E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}^{1/2}\lesssim 2^{-k(d-1-\frac{d+1}{p})}\|f\|_{L^\infty([-1,1]^{d-1})}\,.$$

Note that there are $\mathcal{O}(2^{k(d-1)})$ cubes in $\mathcal{C}_k$, so

$$\|\sum_{\Omega\in\mathcal{C}_k}E_{\Omega_1}fE_{\Omega_2}f\|_{p/2}\lesssim\begin{cases}(2^{k(d-1)}\,2^{-kp(d-1-\frac{d+1}{p})})^{\frac{2}{p}}\|f\|_{L^\infty([-1,1]^{d-1})}^2\,, & d=3\\ (2^k\,2^{-k\frac{2p}{p-2}(1-\frac{3}{p})})^{\frac{p-2}{p}}\|f\|_{L^\infty([-1,1]^{d-1})}^2\,, & d=2\,.\end{cases}$$

In both cases the upper bound is $\mathcal{O}(2^{-k\varepsilon_p}\|f\|_{L^\infty([-1,1]^{d-1})}^2)$ for some $\varepsilon_p>0$ (since $p>2d/(d-1)$). Combining this with (7.18) finishes the proof. □

7.6. **Two-dimensional Kakeya theorems.** We follow Tao [179, Lecture 6].

7.7. **Multilinear restriction.** We follow Bennett–Carbery–Tao [7] and the notes of Hickman and Vitturi [107, Lecture 3, Sections 2-5].

Recall that we have seen in the beginning of this section that the presence of *curvature* of a single (sub)manifold was crucial in the linear restriction problem, whereas *transversality* between two submanifolds became important in the bilinear world. One of the puzzling features of bilinear problems is, however, that they seem to confuse the role played by curvature in higher dimensions. For instance, it is known that the bilinear restriction theories for the cone and the paraboloid are almost identical, whereas the linear theory for these surfaces is certainly not. Moreover, simple heuristics suggest that the optimal k-linear restriction theory requires at least $d-k$ non-vanishing curvatures, but that further curvature assumptions have no further effect. For this reason, it seems natural to consider a d-linear setup in d dimensions since then one does not expect to require any curvature hypotheses. We are therefore seeking inequalities of the form

$$\left\|\prod_{j=1}^d(g_jd\sigma_j)^\vee\right\|_{L^{q/d}(\mathbb{R}^d)}\lesssim\prod_{j=1}^d\|g_j\|_{L^p(S_j)}\quad\text{for all } q\geq 2d/(d-1)\text{ and } p'\leq q(d-1)/d$$

for hypersurfaces $\{S_j\}_{j=1}^d$ endowed with associated smooth measures $\{\sigma_j\}_{j=1}^d$, respectively, whenever the S_j are "sufficiently separated" in the sense of (7.6). In fact, by multilinear interpolation (see, e.g., Bergh–Löfström [8]), and Hölder's inequality, it would suffice to prove the endpoint case $p = 2$ and $q = 2d/(d-1)$, i.e.,

$$\left\| \prod_{j=1}^{d} (g_j d\sigma_j)^\vee \right\|_{L^{2/(d-1)}(\mathbb{R}^d)} \lesssim \prod_{j=1}^{d} \|g_j\|_{L^2(S_j)} .$$

Remarkably, this conjecture was almost completely resolved by Bennett–Carbery–Tao [7] where they proved the above estimate with a subpolynomial loss in the constants.

In the following we adapt the notation that has been used so far to their work. To this end, for $j = 1, ..., d$, let

- $U_j \subseteq \mathbb{R}^{d-1}$ be compact neighborhoods of the origin,
- $\Sigma_j : U_j \to \mathbb{R}^d$ be smooth parameterizations of the $(d-1)$-dimensional manifolds S_j of $\mathbb{R}^d$, and
- $(\mathcal{E}_j g)(x) := \int_{U_j} \mathrm{e}^{2\pi i x \cdot \Sigma_j(\xi)} g(\xi)\, d\xi$ for $x \in \mathbb{R}^d$ be the associated extension operators.

The analog of the bilinear transversality condition will essentially amount to requiring that the normals to the submanifolds parameterized by the Σ_j's span all points of the parameter space. In order to express this in an appropriately uniform manner, we make the following

Definition 7.22. For each $1 \le j \le d$ let Y_j be the $(d-1)$-form

$$Y_j(\xi) := \bigwedge_{k=1}^{d-1} \frac{\partial}{\partial \xi_k} \Sigma_j(\xi) , \quad \xi \in U_j .$$

By duality, the Y_j can be viewed as vector fields on U_j. We will not impose any curvature conditions (in particular, we permit the vector fields Y_j to be constant), but we will impose the following

Assumption 7.23. *Let $A, \nu > 0$ be given. Then the following assertions hold.*

(1) The manifolds S_j obey the "transversality" (or "spanning") condition

$$\det(Y_1(\xi^{(1)}), ..., Y_d(\xi^{(d)})) \ge \nu \quad \textit{for all } \xi^{(1)} \in U_1, ..., \xi^{(d)} \in U_d . \tag{7.19}$$

(2) The maps (parameterizations) Σ_j obey the smoothness condition

$$\|\Sigma_j\|_{C^2(U_j)} \le A \quad \textit{for all } j = 1, ..., d . \tag{7.20}$$

Remarks 7.24. (1) If U_j is sufficiently small, then $\mathcal{E}_j g_j = \widehat{G_j d\sigma_j}$ where $G_j : \Sigma_j(U_j) \to \mathbb{C}$ is the "normalized lift" of g_j, i.e., $G_j(\Sigma_j(\xi)) = |Y_j(\xi)|^{-1} g_j(\xi)$ for $\xi \in U_j$, and $d\sigma_j$ is the induced Lebesgue measure on $\Sigma_j(U_j)$.

(2) Using a partition of unity and an appropriate affine transformation, we can assume $\nu \sim 1$ and that for each $j = 1, ..., d$, the manifold $\Sigma_j(U_j)$ is contained in a sufficiently small neighborhood of the j-th standard basis vector $e_j \in \mathbb{R}^d$.

Observe that, whenever the Σ_j are *linear*, then, by an application of Plancherel's theorem, the conjectured multilinear estimate is equivalent to the Loomis–Whitney inequality [134]. Namely, let $\pi_j : \mathbb{R}^d \to \mathbb{R}^{d-1}$ denote the projection onto the hyperplane $e_j^\perp$ (where $x_j = 0$), i.e., $\pi_j(x) = (x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$, then

$$\int_{\mathbb{R}^d} f_1(\pi_1(x)) \cdots f_d(\pi_d(x))\, dx \le \|f_1\|_{d-1} \cdots \|f_d\|_{d-1} \quad \text{for all } f_j \in L^{d-1}(\mathbb{R}^{d-1}) \tag{7.21}$$

which is sometimes also written as

$$\| \prod_{j=1}^{d} f_j \circ \pi_j \|_{L^{1/(d-1)}(\mathbb{R}^d)} \leq \prod_{j=1}^{d} \|f_j\|_{L^1(e_j^\perp)} \,.$$

For now, let us merely observe that in view of this inequality, we can view multilinear restriction as a certain (rather oscillatory) generalization of the Loomis–Whitney inequality. We will reencounter this inequality in some moments when we will be discussing the multilinear analog of the Kakeya conjecture where the nature of this generalization will become clearer. Let us for now close this subsection with the main result.

Theorem 7.25 (Near-optimal multilinear restriction (Bennett–Carbery–Tao [7, Theorem 1.16]))**.** *Let Assumption 7.23 hold. Then for each $\varepsilon > 0$, $q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$, there exists a constant $C = C(A, \nu, \varepsilon, d, p, q) > 0$ such that*

$$\left\| \prod_{j=1}^{d} \mathcal{E}_j g_j \right\|_{L^{q/d}(B_0(R))} \leq C R^{\varepsilon} \prod_{j=1}^{d} \|g_j\|_{L^p(U_j)} \tag{7.22}$$

holds for all $g_j \in L^p(U_j)$, $j = 1, ..., d$, and all $R \geq 1$.

Naturally, the question arises whether this theorem has any consequences for the linear problem. Unfortunately, the transversality hypotheses make it difficult to apply multilinear restriction estimates directly to obtain new linear estimates in dimensions $d > 2$. After some years however, Bourgain and Guth [31] introduced the so-called ℓ^2-decoupling which allows one to use Theorem 7.25 to obtain improved partial results on the restriction conjecture in higher dimensions. This technique and its applications will be discussed in detail in Section 21.

In Subsection 7.10 we will see that this theorem is equivalent to the so-called multilinear Kakeya conjecture that we will discuss (and prove!) in the next subsections.

7.8. **Multilinear Kakeya.** We follow Bennett–Carbery–Tao [7] and the notes of Hickman and Vitturi [107, Lecture 3, Sections 2-5]. See also Guth [100, 101] (in particular the short proof of the non-optimal result.)

It is well known (and it will be discussed in Subsection 15.1) that the linear restriction conjecture implies the linear Kakeya conjecture (Conjecture 15.1). To state it precisely, let us introduce the following notation that will also be used in Section 15 later. Let $0 < \delta \ll 1$, $\omega \in \mathbb{S}^{d-1}$, and $a \in \mathbb{R}^d$. Then we define a *δ-tube* to be any rectangular (or cylindrical) box $T_\omega^\delta(a)$, or short, T, in $\mathbb{R}^d$ with $d-1$ side lengths δ (or diameter 2δ) and one side length 1 which is oriented in the direction ω. By $\mathbb{T}$ we denote an arbitrary collection of such δ-tubes whose orientations form a maximal δ-separated subset of $\mathbb{S}^{d-1}$. The cardinality of $\mathbb{T}$ is denoted by $\#\mathbb{T}$. Then the maximal Kakeya conjecture says that for any $\varepsilon > 0$ and $d/(d-1) < q \leq \infty$, there exists a constant C independent of δ such that

$$\left\| \sum_{T \in \mathbb{T}} \mathbf{1}_T \right\|_{L^q(\mathbb{R}^d)} \leq C \delta^{(d-1)/q} (\#\mathbb{T})^{1-1/(q(d-1))} \,.$$

We emphasize that the "separation condition" on each of the δ-tubes is crucial in this linear problem.

By a straightforward adaptions of the arguments given in Subsection 15.1, one sees that the multilinear restriction conjecture implies the corresponding multilinear Kakeya-type conjecture that we will describe now. Suppose $\mathbb{T}_1, ..., \mathbb{T}_d$ are families of δ-tubes in $\mathbb{R}^d$ where we now allow (!) the tubes within the same family $\mathbb{T}_j$ to be parallel (in contrast to the linear problem). However, we assume that, for each $j = 1, ..., d$, the tubes in $\mathbb{T}_j$ must point in directions belonging to a fixed spherical cap, say $S_j = \{\omega \in \mathbb{S}^{d-1} : \ |1 - \omega \cdot e_j| \le C^{-1}\}$ for some large $C > 0$, centered at e_j. In this case, we say that *the family* $\mathbb{T}_j$ *is transversal.* (The vectors e_j may be replaced by any fixed linearly independent set of vectors in $\mathbb{R}^d$ here, as affine invariance considerations reveal.)

Theorem 7.26 (Near-optimal multilinear Kakeya (Bennett–Carbery–Tao [7, Theorem 1.15])). *If $d/(d-1) < q \le \infty$, then there exists a constant $C > 0$ which is independent of δ and the transversal families of tubes $\mathbb{T}_1, ...\mathbb{T}_d$, such that*

$$\Big\| \prod_{j=1}^{d} \Big(\sum_{T_j \in \mathbb{T}_j} \mathbf{1}_{T_j} \Big) \Big\|_{L^{q/d}(\mathbb{R}^d)} \le C \prod_{j=1}^{d} (\delta^{d/q} \# \mathbb{T}_j) \,. \tag{7.23}$$

Furthermore, for each $\varepsilon > 0$ there is a similarly uniform constant $C > 0$ for which

$$\Big\| \prod_{j=1}^{d} \Big(\sum_{T_j \in \mathbb{T}_j} \mathbf{1}_{T_j} \Big) \Big\|_{L^{1/(d-1)}(B_0(1))} \le C \delta^{-\varepsilon} \prod_{j=1}^{d} (\delta^{d-1} \# \mathbb{T}_j) \,. \tag{7.24}$$

Remarks 7.27. (1) Since the case $q = \infty$ is trivially true, (7.23) is equivalent, via Hölder, to the endpoint case $q = d/(d-1)$. In contrast to the linear setting, there is no obvious counterexample prohibiting this claim holding at the endpoint $q = d/(d-1)$, and indeed in the $d = 2$ case it is easy to verify this endpoint estimate. In fact, as we will present next, Guth [100] did eventually obtain the endpoint result.

(2) By contrast with similar statements at lower levels of multilinearity, each family $\mathbb{T}_j$ is permitted to contain parallel tubes, and in fact, even arbitrary repetitions of tubes.

(3) The decision to formulate (7.23) in terms of $\delta \times \cdots \delta \times 1$ tubes is largely for historical reasons. However, just by scaling, it is easily seen that the above estimate is equivalent to

$$\Big\| \prod_{j=1}^{d} \Big(\sum_{\tilde{T}_j \in \tilde{\mathbb{T}}_j} \mathbf{1}_{\tilde{T}_j} \Big) \Big\|_{L^{q/d}(\mathbb{R}^d)} \le C \prod_{j=1}^{d} (\# \tilde{\mathbb{T}}_j) \,,$$

where the collections $\tilde{\mathbb{T}}_j$ consist of tubes of width 1 and arbitrary (possibly infinite) length where, of course, still the appropriate transversality condition is imposed on the families $\tilde{\mathbb{T}}_1, ..., \tilde{\mathbb{T}}_d$.

(4) Note that the extreme case of (7.23) when the collections of rectangles are 1-transverse corresponds (by Hadamard's inequality) precisely to the situation when all the rectangles in $\mathbb{T}_j$ are oriented in the same direction e_j. Under these hypothesis, (7.23) is a consequence of the Loomis–Whitney inequality (7.21). Put differently, the multilinear Kakeya estimate is a generalization of the Loomis–Whitney inequality. The geometric nature of this generalization is of course much more transparent than in the multilinear restriction problem.

(4) As opposed to the linear case, the multilinear Kakeya theorem does not imply something on the dimension of Besicovitch sets, although there is a connection to the joints problem, see Bennett–Carbery–Tao [7, §7].

(5) Bennett–Carbery–Tao [7, §6] also derived a natural variable coefficient extension of their results.

(6) Although, we will not review their proof here, let us summarize their strategy. First, one observes that if each $T_j \in \mathbb{T}_j$ is centered at the origin (for all $j = 1, ..., d$), then, the two sides of (7.23) are trivially comparable. This observation leads to the suggestion that such configurations of tubes might actually be *extremal* for the left side of (7.23). For analytic reasons, in pursuing this idea it seemed natural to replace the rough indicator functions by Gaussians of the form $\mathrm{e}^{-\langle x-v, A(x-v)\rangle}$ for an appropriate positive definite $d \times d$ matrix A and vectors $v \in \mathbb{R}^d$. Using these Gaussians as "smoothed cutoff functions", they give a novel proof of the Loomis–Whitney inequality in §3. Afterwards, they perturb the inequality in §4; as a corollary of this perturbed

inequality, they obtain the multilinear Kakeya conjecture up to the endpoint (and a "weak" form of the multilinear restriction conjecture). □

Theorem 7.28 (Weighted multilinear Kakeya)**.** *Assume that the assumptions of Theorem 7.26 hold. For each $T_j \in \mathbb{T}_j$ let $w_{T_j} \geq 0$ be a weight and define the simple functions*

$$g_j := \sum_{T_j \in \mathbb{T}_j} w_{T_j} \mathbf{1}_{T_j} \,.$$

Then, one has the similar estimate

$$\left\| \prod_{j=1}^d g_j \right\|_{L^{1/(d-1)}(B_0(1))} \leq C\delta^{-\varepsilon} \prod_{j=1}^d (\delta^{d-1} \sum_{T_j \in \mathbb{T}_j} w_{T_j}) \,. \tag{7.25}$$

Proof. If $w_{T_j} \in \mathbb{N}$ for all $T_j \in \mathbb{T}_j$, then the result is a consequence of the original multilinear Kakeya inequality (7.23) by including repeats of the tubes in the collections. The estimate for rational weights follows by rescaling and for reals by continuity. □

Theorem 7.29 (Endpoint multilinear Kakeya (Guth [100]))**.** *Formula* (7.24) *holds without the subpolynomial loss $\delta^{-\varepsilon}$. Moreover, the dependence of the transversality constant ν is given by $\nu^{-1/(d-1)}$.*

Theorem 7.30 (Simple multilinear Kakeya (Guth [101]))**.** *Suppose that $\ell_{j,a}$ are lines in $\mathbb{R}^d$ where $j = 1, ..., d$ and $a = 1, ..., N_j$. Let $\tilde{T}_{j,a}$ be the 1-neighborhood of $\ell_{j,a}$. Suppose that $S_j \subseteq \mathbb{S}^{d-1}$ is a spherical cap and that the lines $\ell_{j,a}$ lie in S_j. Suppose that for any vectors $v_j \in S_j$, we have the transversality condition $|v_1 \wedge \cdots \wedge v_d| \geq \nu$.*

Let Q_S denote any cube of side length S. Then for any $\varepsilon > 0$ and any $S \geq 1$, one has

$$\int_{Q_S} \prod_{j=1}^d \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq C_\varepsilon \nu^{-\mathcal{O}(1)} S^\varepsilon \prod_{j=1}^d N_j^{\frac{1}{d-1}} \,, \tag{7.26}$$

where $\nu^{-\mathcal{O}(1)}$ means that the dependence on the transversality constant ν is polynomial.

Moreover, we have the following weighted analog similar to Theorem 7.28. For each $T_{j,a} \in \mathbb{T}_j$ let $w_{j,a} \geq 0$ be a weight and define the simple functions

$$g_j := \sum_{T_j \in \mathbb{T}_j} w_{j,a} \mathbf{1}_{T_{j,a}} \,.$$

Then, with the above notation,

$$\int_{Q_S} \prod_{j=1}^d g_j^{\frac{1}{d-1}} \leq C_\varepsilon \nu^{-\mathcal{O}(1)} S^\varepsilon \prod_{j=1}^d \left(\sum_a w_{j,a} \right)^{\frac{1}{d-1}} \,, \tag{7.27}$$

holds.

Remarkably, we will find next that the multilinear restriction and Kakeya theorems are essentially *equivalent*. This equivalence follows from multilinearizing a well known induction-on-scales argument of Bourgain [15] (see also Tao–Vargas–Vega [188] for this argument in the bilinear setting). Before we study this equivalence in detail, we proceed with a review of Guth's simple proof of Theorem 7.30.

7.9. **Guth's simple proof of Theorem 7.30.** The main goal of this section will be to prove the following theorem. Theorem 7.30 will follow from it and the ensuing observation.

Theorem 7.31. *Suppose that $\ell_{j,a}$ are lines in $\mathbb{R}^d$ where $j = 1, ..., d$ and $a = 1, ..., N_j$. Let $\tilde{T}_{j,a}$ be the 1-neighborhood of $\ell_{j,a}$. Suppose that the lines $\ell_{j,a}$ makes an angle of at most $(10d)^{-1}$ with the e_j-axis.*

Let Q_S denote any cube of side length S. Then for any $\varepsilon > 0$ and any $S \geq 1$, one has

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq C_\varepsilon S^\varepsilon \prod_{j=1}^{d} N_j^{\frac{1}{d-1}}, \tag{7.28}$$

The proof is split into three steps.

(1) Reduction to almost axis-parallel tubes
(2) Analyzing the case of exactly axis-parallel tubes using the Loomis–Whitney inequality (7.21)
(3) Perturbation of the Loomis–Whitney inequality and multiscale analysis

7.9.1. *Reduction to nearly axis parallel tubes.* The first observation in Bennett et al [7] is that it suffices to consider collections $\mathbb{T}_j$ of tubes which are almost parallel to each other. In fact, Theorem 7.31 will follow from

Proposition 7.32. *For every $\varepsilon > 0$, there is some $\delta > 0$ such that the following holds. Suppose that $\ell_{j,a}$ are lines in $\mathbb{R}^d$, and that each line $\ell_{j,a}$ makes an angle of at most δ with the e_j-axis. Then for any $S \geq 1$ and any cube Q_S of side length S, we have*

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq C_\varepsilon S^\varepsilon \prod_{j=1}^{d} N_j^{\frac{1}{d-1}}. \tag{7.29}$$

We will use this to prove Theorems 7.31 and 7.30.

Proof of Theorem 7.31 assuming Proposition 7.32. Let $S_j \subseteq \mathbb{S}^{d-1}$ be a spherical cap around e_j of radius, say $(10d)^{-1}$. By the hypothesis of Theorem 7.31, every line $\ell_{j,a}$ has a direction belonging to S_j. Now, for a given $\varepsilon > 0$, we pick a δ as in Proposition 7.32. We subdivide S_j now into smaller caps $S_{j,\beta}$ of radius $\delta/10$, i.e., S_j can be covered by roughly $\delta^{-1} \lesssim_\varepsilon 1$ caps $S_{j,\beta}$. Let us abuse notation and write "$\ell_{a,j} \in S_{j,\beta}$", whenever the direction of $\ell_{a,j}$ belongs to $S_{j,\beta}$. Since the number of caps is $\lesssim_\varepsilon 1$, we have

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \lesssim_\varepsilon \sum_{\beta_1,\ldots,\beta_d} \int_{Q_s} \prod_{j=1}^{d} \left(\sum_{\ell_{j,a} \in S_{j,\beta}} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}}.$$

We claim that each β-summand on the right side is controlled by Proposition7.32. Clearly, this is the case when β_j is such that S_{j,β_j} contains e_j. Otherwise, we perform a linear change of variables such that the center of $S_{j,\beta}$ is mapped to e_j. Since the angle between the $\ell_{j,a}$ and e_j is at most $(10d)^{-1}$, the involved Jacobian is at most const^d. In any case, the integral in the new coordinates is again controlled using Proposition 7.32. □

Proof of Theorem 7.30 assuming Proposition 7.32. We again cover S_j by caps $S_{j,\beta}$ of a small radius ρ. As long as $\rho \leq \nu/(100d)$, we can guarantee that $|v_1 \wedge \cdots \wedge v_d| \geq \nu/2$ for all $v_j \in S_{j,\beta}$. We pick a sequence of caps $S_{1,\beta_1}, ..., S_{d,\beta_d}$ and change coordinates so that the center of the cap S_{j,β_j} is mapped to the coordinate vector e_j. The distortion of lengths and volumes caused by this coordinate change is $\mathcal{O}(\nu^{-1})$. So, we may apply Proposition 7.32 in these new coordinates.

If $\rho = \rho(\varepsilon)$ is small enough, the image of $S_{j,\beta}$ is contained in a cap of radius $\delta = \delta(\varepsilon)$ as in Proposition 7.32 – and this gives the desired estimate with error of order $C_\varepsilon \mathcal{O}(\nu^{-1}) S^\varepsilon$. Finally, we sum over $C_\varepsilon \mathcal{O}(\nu^{-1})$ with different choices of $S_{1,\beta_1}, ..., S_{d,\beta_d}$. □

7.9.2. *The axis parallel case (Loomis–Whitney).* As we have remarked after Theorem 7.26, the case where all $\ell_{j,a}$ are parallel to the e_j-axis follows from the Loomis–Whitney inequality (7.21) in the form

$$\int_{\mathbb{R}^d} \prod_{j=1}^d f_j(\pi_j(x))^{\frac{1}{d-1}} \leq \prod_{j=1}^d \|f_j\|_{L^1(\mathbb{R}^{d-1})}^{\frac{1}{d-1}} .$$

In fact, if the line $\ell_{j,a}$ is parallel to the e_j-axis, then it can be defined by the point $\pi_j(x) = y_a \in \mathbb{R}^{d-1}$ where it intersects the plane $x_j = 0$, see the figure below.

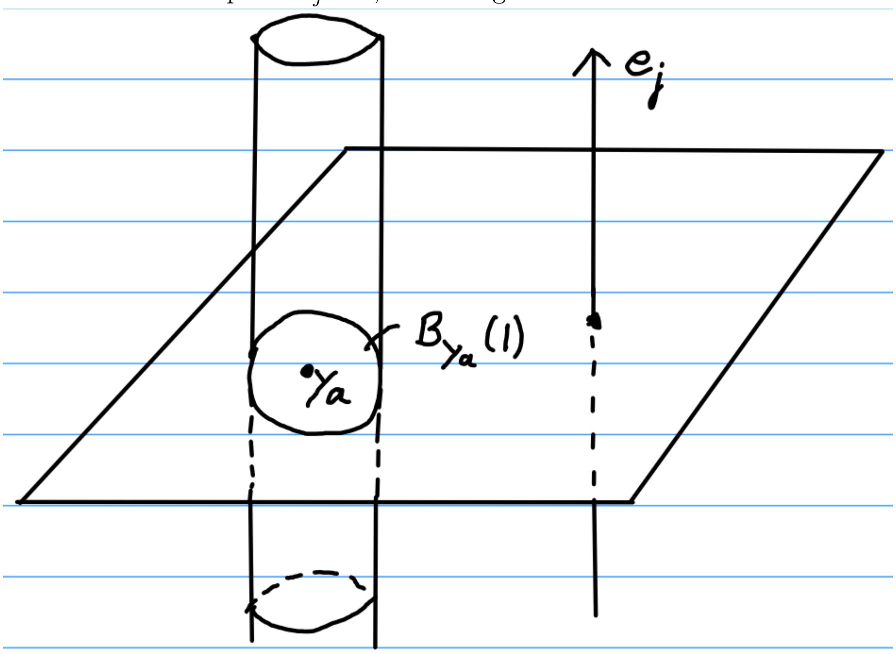

Figure 2

Then

$$\sum_a \mathbf{1}_{T_{j,a}}(x) = \sum_a \mathbf{1}_{B_{y_a}(1)}(\pi_j(x)) .$$

Applying the Loomis–Whitney inequality with

$$f_j = \sum_a \mathbf{1}_{B_{y_a}(1)}(\pi_j(x))$$

with $\|f_j\|_{L^1(\mathbb{R}^{d-1})} = |\mathbb{S}^{d-1}| N_j$, we obtain

$$\int_{\mathbb{R}^d} \prod_{j=1}^d \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} = \int_{\mathbb{R}^d} \prod_{j=1}^d f_j(\pi_j(x))^{\frac{1}{d-1}} \leq \prod_{j=1}^d \|f_j\|_{L^1(\mathbb{R}^{d-1})} \sim \prod_{j=1}^d N_j^{\frac{1}{d-1}} .$$

7.9.3. *The multiscale argument.* In the previous subsubsection we saw how to prove Proposition 7.32 in the case where all tubes are parallel to each other. We will now have to understand and control the impact of slightly tilting them with tilting angle at most $\delta = \delta(\varepsilon)$ for a given fixed ε. The main idea is the following. Instead of trying to prove the desired estimate immediately on the scale S, we will first study a smaller scale, say δ^{-1}. Then, we will jump to the larger scale δ^{-2} using the Loomis–Whitney inequality and continue this procedure until we arrive at the desired scale S.

To set up the argument (and also generalize the lemma a bit), we introduce the one-parameter family of tubes of *variable thickness* $T_{j,a,W}$ which are W-neighborhoods (cylindrical or rectangular) of the line $\ell_{j,a}$.

The following lemma is crucial to get the inductive step from scale δ^{-1} to scale δ^{-2} running.

Lemma 7.33. *Suppose the lines $\ell_{j,a}$ make an angle of at most δ from the e_j-axis. Let $T_{j,a,W}$ be as before and introduce*

$$f_{j,W} := \sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a,W}} .$$

If $S \geq W/\delta$ and Q_S is any cube of sidelength S, then

$$\int_{Q_S} \prod_{j=1}^{d} f_{j,W}^{1/(d-1)}\, dx \leq C_d \delta^d \int_{Q_S} \prod_{j=1}^{d} f_{j,W/\delta}^{1/(d-1)}\, dx .$$

Proof. Since $S \geq W/\delta$, we may divide Q_S into subcubes Q whose side length belongs to $[\frac{W/\delta}{20d}, \frac{W/\delta}{10d}]$. Thus, it suffices to prove for each such cube

$$\int_{Q} \prod_{j=1}^{d} f_{j,W}^{1/(d-1)}\, dx \leq C_d \delta^d \int_{Q} \prod_{j=1}^{d} f_{j,W/\delta}^{1/(d-1)}\, dx .$$

Since the side length of Q is $\leq \frac{W/\delta}{10d}$, one can find an axis-parallel tube $\tilde{T}_{j,a,\tilde{W}}$ of twice the thickness, i.e., $\tilde{W} = 2W$, see also the figure below.

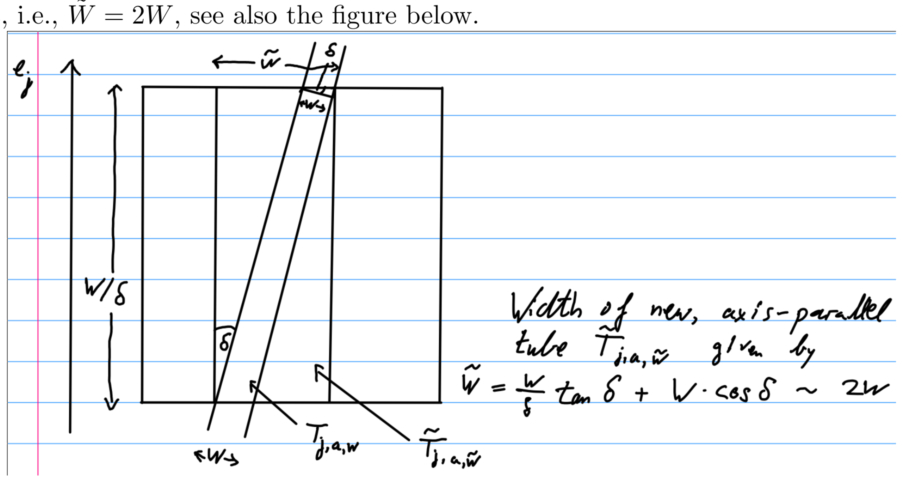

Figure 3

Therefore, we have $\mathbf{1}_{T_{a,j,W}}(x) \leq \mathbf{1}_{\tilde{T}_{a,j,2W}}(x)$ for all $x \in Q$ and may estimate

$$\int_Q \prod_{j=1}^{d} f_{j,W}^{1/(d-1)} \leq \int_Q \prod_{j=1}^{d} \left(\sum_a \mathbf{1}_{\tilde{T}_{a,j,2W}} \right)^{\frac{1}{d-1}} \lesssim \prod_{j=1}^{d} \left(N_j(Q)^{\frac{1}{d-1}} \cdot W^{\frac{d-1}{d-1}} \right) = W^d \prod_{j=1}^{d} N_j(Q)^{\frac{1}{d-1}} .$$

Here, we used Loomis–Whitney in the second inequality (like in the previous step) and denoted the number of tubes $T_{j,a,W}$ that intersect Q by $N_j(Q)$.

Now, since the side length of Q is $\leq \frac{W/\delta}{10d}$, its diameter is $\leq \frac{W/\delta}{10\sqrt{d}} \leq \frac{W/\delta}{10}$. Thus, if $T_{j,a,W}$ intersects Q, then certainly $\mathbf{1}_{T_{j,a,W/\delta}}(x) = 1$ for all $x \in Q$ and hence $N_j(Q) \leq \sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a,W/\delta}}(x)$ for all $x \in Q$. Using this bound and that $|Q| \sim (W/\delta)^d$, we obtain

$$W^d \prod_{j=1}^{d} N_j(Q)^{\frac{1}{d-1}} \lesssim \delta^d |Q| \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a,W/\delta}}(x) \right)^{\frac{1}{d-1}} \sim \delta^d \int_Q \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a,W/\delta}}(x) \right)^{\frac{1}{d-1}},$$

thereby establishing the claim. □

We are now ready to give the

Proof of Theorem 7.32. Suppose first $S = \delta^{-M}$. Using Lemma 7.33 repeatedly, we get

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} = \int_{Q_S} \prod_{j=1}^{d} f_{j,1}^{\frac{1}{d-1}} \leq C_d^M \delta^{d \cdot M} \int_{Q_S} \prod_{j=1}^{d} f_{j,\delta^{-M}}^{\frac{1}{d-1}}$$

with C_d from the assertion of that lemma. Since $f_{j,\delta^{-M}}(x) \leq N_j$ for all x, we can further estimate

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq C_d^M \delta^{d \cdot M} \prod_{j=1}^{d} N_j^{\frac{1}{d-1}} \int_{Q_S} = C_d^M \prod_{j=1}^{d} N_j^{\frac{1}{d-1}}.$$

Since $S = \delta^{-M}$, we have $M = \frac{\log S}{\log(\delta^{-1})}$ and therefore $C_d^M = S^{\frac{\log C_d}{\log(\delta^{-1})}}$. Now, for given ε, we chose $\delta = \delta(\varepsilon)$ so small that $\frac{\log C_d}{\log(\delta^{-1})} \leq \varepsilon$. Thus, for $S = \delta^{-M}$, the above estimate reads

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq S^\varepsilon \prod_{j=1}^{d} N_j^{\frac{1}{d-1}}$$

when $S = \delta^{-M}$. Now, for an arbitrary $S \geq 1$, we can find $M \in \mathbb{N}_0$ so that Q_S can be covered by $C_{\delta(\varepsilon)}$ cubes of side length δ^{-M}. But then we can use the above estimate for each such subcube and obtain

$$\int_{Q_S} \prod_{j=1}^{d} \left(\sum_{a=1}^{N_j} \mathbf{1}_{T_{j,a}} \right)^{\frac{1}{d-1}} \leq C_\varepsilon S^\varepsilon \prod_{j=1}^{d} N_j^{\frac{1}{d-1}}.$$

This concludes the proof of Theorem 7.32. □

7.10. **Multilinear restriction ⇔ multilinear Kakeya.** We follow Bennett–Carbery–Tao [7, §2].

Notation. Recall that we introduced for $\alpha \geq 0$, $q \geq 2d/(d-1)$, and $p' \leq q(d-1)/d$ the notation

$$\mathcal{R}^*(p \times ... \times p \to q; \alpha)$$

to denote the multilinear restriction estimate

$$\| \prod_{j=1}^{d} \mathcal{E}_j g_j \|_{L^{q/d}(B_R(0))} \leq C R^\alpha \prod_{j=1}^{d} \|g_j\|_{L^p(U_j)},$$

for some $C = C(A, \nu, \alpha, d, p, q)$, for all $g_j \in L^p(U_j)$, $j = 1, ..., d$, and all $R \geq 1$. Similarly, for $d/(d-1) \leq q \leq \infty$, we use

$$\mathcal{K}^*(1 \times ... \times 1 \to q; \varepsilon)$$

to denote the multilinear Kakeya estimate

$$\| \prod_{j=1}^{q} (\sum_{T_j \in \mathbb{T}_j} \mathbf{1}_{T_j}) \|_{q/d} \leq C\delta^{-\varepsilon} \prod_{j=1}^{d} (\delta^{d/q} \# \mathbb{T}_j) \tag{7.30}$$

for some $C = C(\varepsilon, d, q)$, for all transversal collections of families of δ-tubes in $\mathbb{R}^d$, and all $0 < \delta \leq 1$. Recall once more, that (7.30) is equivalent (by standard density arguments in suitable weak topologies), to the superficially stronger inequality

$$\| \prod_{j=1}^{d} (\sum_{T_j \in \mathbb{T}_j} \mathbf{1}_{T_j} * \mu_{T_j}) \|_{L^{q/d}(\mathbb{R}^d)} \leq C\delta^{-\varepsilon} \prod_{j=1}^{d} (\delta^{d/q} \sum_{T_j \in \mathbb{T}_j} \|\mu_{T_j}\|) \tag{7.31}$$

for all finite measures μ_{T_j} (with $T_j \in \mathbb{T}_j$ and $j = 1, ..., d$) on $\mathbb{R}^d$.

With this notation, Theorem 7.26 is equivalent to the statement $\mathcal{K}^*(1 \times ... \times 1 \to q; 0)$ for all $d/(d-1) \leq q \leq \infty$, and $\mathcal{K}^*(1 \times ... \times 1 \to d/(d-1); \varepsilon)$ for all $\varepsilon > 0$. Similarly, Theorem 7.25 is equivalent to $\mathcal{R}^*(2 \times ... \times 2 \to 2d(d-1); \varepsilon)$ for all $\varepsilon > 0$.

7.10.1. *Multilinear restriction ⇒ multilinear Kakeya.* As we have already outlined (see also Proposition 15.4), a standard randomization argument allows one to deduce the multilinear Kakeya conjecture from the multilinear restriction conjecture. In the localized setting, this of course continues to be true, i.e., for any $\alpha \geq 0$, we have

$$\mathcal{R}^*(2 \times ... \times 2 \to \frac{2d}{d-1}; \alpha) \Rightarrow \mathcal{K}^*(1 \times ... \times 1 \to \frac{d}{d-1}; 2\alpha)\,. \tag{7.32}$$

7.10.2. *Multilinear Kakeya ⇒ multilinear restriction.* Multilinearizing a well known bootstrapping argument of Bourgain [15] (again, see Tao–Vargas–Vega [188] in the bilinear setting), we shall obtain the following reverse mechanism.

Proposition 7.34. *For all $\alpha, \varepsilon \geq 0$ and $2d/(d-1) \leq q \leq \infty$, we have*

$$\mathcal{R}^*(2 \times ... \times 2 \to q; \alpha) \text{ and } \mathcal{K}^*(1 \times ... \times 1 \to \frac{q}{2}; \varepsilon) \Rightarrow \mathcal{R}^*(2 \times ... \times 2 \to q; \frac{\alpha}{2} + \frac{\varepsilon}{4})\,.$$

Remark 7.35. Note that there are minor flaws in the proofs of this proposition in Bennett–Carbery–Tao [7, Proposition 2.1] and in Bennett [6, Proposition 4.8]. (Formula (14) in [7] can only hold, when the $L^2(A_j^R)$ norm on the right side of the estimate is replaced by the $L^2(A_j^{\sqrt{R}})$ norm. A similar flaw occurs in [6].) This flaw is however not grave, as $A_j^{\sqrt{R}}$ can still be covered by $R^{-1/2} \times ... \times R^{-1/2} \times R^{-1}$ discs (there are now $\mathcal{O}(R^{1/2})$ more discs in the argument as in these works) as they are merely used to perform a partition of unity of $f_j * \varphi^x_{R^{1/2}}$. In any case, a correct version of the proof appears Lecture 1 (Proposition 36) in Tao's notes [187].

Using elementary estimate, one easily verifies $\mathcal{R}^*(2 \times ... \times 2 \to 2d/(d-1); \alpha)$ for very large α. For instance, noting that $|B_0(R)| = c_d R^d$, one has

$$\| \prod_{j=1}^{d} \mathcal{E}_j g_j \|_{L^{2/(d-1)}(B_0(R))} \leq c_d R^{d(d-1)/2} \prod_{j=1}^{d} \|\mathcal{E}_j g_j\|_\infty \leq c_d R^{d(d-1)/2} \prod_{j=1}^{d} \|g_j\|_{L^1(U_j)}\,,$$

which, by Cauchy–Schwarz, yields $\mathcal{R}^*(2 \times ... \times 2 \to 2d/(d-1); d(d-1)/2)$. In the presence of appropriately favorable Kakeya estimates this large value of α may then be reduced by a repeated application of the above proposition. In particular, Proposition 7.34 together with (7.32) (multilinear restriction ⇒ multilinear Kakeya) shows the equivalence

$$\mathcal{R}^*(2 \times ... \times 2 \to \frac{2d}{d-1}; \varepsilon) \Leftrightarrow \mathcal{K}^*(1 \times ... \times 1 \to \frac{d}{d-1}; \varepsilon) \quad \text{for all } \varepsilon > 0\,.$$

Therefore, the multilinear restriction theorem (Theorem 7.26) follows from the multilinear Kakeya theorem (Theorem 7.26). In fact, already Guth's simpler version (Theorem 7.30) is sufficient to prove Theorem 7.26.

The proof of Proposition 7.34 is very similar to that of Tao–Vargas–Vega [188, Lemma 4.4], and on a technical level slightly more straightforward. We begin by stating a lemma which, given (1) in Remark 7.24 and the control of $|Y_j|$ implicit in Assumption 7.23 is a standard manifestation of the uncertainty principle. (See Córdoba [49] for the origin of this (see also [51]) and Tao–Vargas–Vega [188, Proposition 4.3] for a proof in the bilinear case which immediately generalizes to the multilinear case.) In effect, this is similar to localized *linear* restriction theory that we discussed in Section 6, especially Lemmas 6.1 and 6.3.

Lemma 7.36. *The multilinear restriction estimate $\mathcal{R}^*(2\times ... \times 2 \to q;\alpha)$ is true if and only if*

$$\| \prod_{j=1}^d \check{f}_j \|_{L^{q/d}(B_0(R))} \leq C R^{\alpha-d/2} \prod_{j=1}^d \|f_j\|_2 \tag{7.33}$$

for all $R\geq 1$ and functions $f_j \in \hat{\mathcal{S}}(\mathbb{R}^d)$ with $\operatorname{supp} f_j \subseteq A_j^R \equiv \mathcal{N}_{1/R}(\Sigma_j(U_j)) := \Sigma_j(U_j)+\mathcal{O}(R^{-1})$ *(an R^{-1}-neighborhood or R^{-1}-annulus) for all $j=1,...,d$.*

We now turn to the proof of Proposition 7.34, where the implicit constants in the $\lesssim$ notation will at most depend on A,ν,d,p,α, and ε.

Proof of Proposition 7.34. The proof is somewhat similar to the one of Lemma 7.33 and uses induction on scales.

Because of the above lemma on the equivalence between global and localized restriction estimates, it suffices to show

$$\| \prod_{j=1}^d \check{f}_j \|_{L^{q/d}(B_0(R))} \lesssim R^{\alpha/2+\varepsilon/4-d/2} \prod_{j=1}^d \|f_j\|_{L^2(A_j^R)}$$

for all f_j with $\operatorname{supp} f_j \subseteq A_j^R$ with $j=1,...,d$. To this end, let $\varphi \in \widehat{C_c^\infty}(\mathbb{R}^d)$ be a bump function adapted to $B_0(C)$ such that $\check{\varphi}(x)\geq 0$ for $x\in B_0(1)$. For $R\geq 1$ and $x\in\mathbb{R}^d$, define the modulated (L^1-normalized) dilate $\varphi^x_{R^{1/2}}(\xi) := \mathrm{e}^{-2\pi i x\cdot\xi} R^{d/2}\varphi(R^{1/2}\xi)$ which is a bump adapted to $B_0(C/R^{1/2})$ in Fourier space, respectively a Schwartz function in physical space which is centered at x, bounded from below on $B_x(C^{-1}R^{1/2})$, and rapidly decaying away from $B_x(C^{-1}R^{1/2})$. By the assumption $\mathcal{R}^*(2\times ... \times 2 \to q;\alpha)$ and the above localization lemma (Lemma 7.36 with R replaced by $\sqrt{R}$ and replacing f_j by the modulate $f_j \mathrm{e}^{2\pi i\langle x,\cdot\rangle}$), we infer

$$\| \prod_{j=1}^d (\varphi^x_{R^{1/2}})^\vee \check{f}_j \|_{L^{q/d}(B_x(R^{1/2}))} \lesssim R^{\alpha/2-d/4} \prod_{j=1}^d \|\varphi^x_{R^{1/2}} * f_j\|_{L^2(A_j^{\sqrt{R}})}$$

for all $x\in\mathbb{R}^d$. Thus, "$L^{q/d}$-averaging" this inequality over $x\in B_0(R)$ (i.e., taking both sides to the power q/d, integrating over $x\in B_0(R)$, and taking everything to the power d/q) yields (using $\int_{|x|\leq R} \mathbf{1}_{|x-y|\leq\sqrt{R}}\,dx \gtrsim R^{d/2}\mathbf{1}_{|y|\leq R}$)

$$\| \prod_{j=1}^d \check{f}_j \|_{L^{q/d}(B_0(R))} \lesssim R^{\alpha/2-d/4} \left[R^{-d/2} \int_{B_0(R)} \left(\prod_{j=1}^d \|f_j * \varphi^x_{R^{1/2}}\|^2_{L^2(A_j^{\sqrt{R}})} \right)^{q/(2d)} dx \right]^{d/q}.$$

In what follows, we shall show that the $[...]^{d/q}$-term is bounded by $R^{\varepsilon/4}\cdot R^{-d/4}$ where the $R^{\varepsilon/4}$ is just the square root of the constant in the multilinear Kakeya estimate (7.31) with

$\delta = R^{-1/2}$. Now, for each $j = 1, ..., d$, we cover $A_j^{\sqrt{R}}$ (the $R^{-1/2}$-neighborhood around $\Sigma_j(U_j)$) by a boundedly overlapping family of $R^{-1/2} \times ... \times R^{-1/2} \times R^{-1}$-discs $\{D_j\}$ and introduce $f_{j,D_j} := \mathbf{1}_{D_j} f_j$. Since for each j, the supports of the functions $f_{j,D_j} * \varphi^x_{R^{1/2}}$ are only finitely overlapping, we further obtain

$$\| \prod_{j=1}^d \check{f}_j \|_{L^{q/d}(B_0(R))} \lesssim R^{\alpha/2-d/4} \left[R^{-d/2} \int_{B_0(R)} \left(\prod_{j=1}^d \sum_{D_j} \| f_{j,D_j} * \varphi^x_{R^{1/2}} \|^2_{L^2(\mathbb{R}^d)} \right)^{q/(2d)} dx \right]^{d/q} .$$

Applying Plancherel to the right side and using that $(\varphi^x_{R^{1/2}})^\vee$ is rapidly decreasing away from $B_x(\sqrt{R})$, we estimate the right side further from above by a constant times

$$R^{\alpha/2-d/4} \left[R^{-d/2} \int_{B_0(R)} \left(\prod_{j=1}^d \sum_{D_j} \| (f_{j,D_j})^\vee \|^2_{L^2(B_x(R^{1/2}))} \right)^{q/(2d)} dx \right]^{d/q} . \qquad (7.34)$$

For each D_j, let $\psi_{D_j} \in \hat{\mathcal{S}}(\mathbb{R}^d)$ with $\psi_{D_j}(\xi) \sim 1$ for $\xi \in D_j$ and whose Fourier transform satisfies

$$|(\psi_{D_j})^\vee(x+y)| \lesssim R^{-(d+1)/2} \mathbf{1}_{T_j}(x) , \quad x, y \in \mathbb{R}^d \text{ with } |y| \leq R^{1/2} ,$$

where T_j denotes the $R^{1/2} \times ... \times R^{1/2} \times R$-tube (which is dual to the disc D_j) centered at the origin and oriented along the normal of the disc D_j. (Note that we are here using the full $C^2(U_j)$ control given by Assumption 7.23.) Defining $\tilde{f}_{j,D_j} := f_{j,D_j}/\psi_{D_j}$, we see that f_{j,D_j} and $\tilde{f}_{j,D_j}$ are pointwise comparable. Now, by Cauchy–Schwarz (write the following convolution like $|(\tilde{f}_{j,D_j})^\vee(z)| \, |(\psi_{D_j})^\vee(z-w)|^{1/2} \cdot |(\psi_{D_j})^\vee(z-w)|^{1/2}$), we may estimate

$$|(f_{j,D_j})^\vee(x+y)|^2 = |(\tilde{f}_{j,D_j})^\vee * (\psi_{D_j})^\vee(x+y)|^2 \lesssim R^{-(d+1)/2} |(\tilde{f}_{j,D_j})^\vee|^2 * \mathbf{1}_{T_j}(x)$$

for all $x, y \in \mathbb{R}^d$ with $|y| \leq R^{1/2}$. Integrating this in y over $|y| \leq R^{1/2}$ yields

$$\| (f_{j,D_j})^\vee \|^2_{L^2(B_x(R^{1/2}))} \lesssim R^{-1/2} |(\tilde{f}_{j,D_j})^\vee|^2 * \mathbf{1}_{T_j}(x) .$$

Plugging this estimate in (7.34) and applying the $R^{-1/2}$-rescaled Kakeya hypothesis $\mathcal{K}^*(1 \times ... \times 1 \to q/2; \varepsilon)$ with $\delta = R^{-1/2}$ (in its equivalent "measure form" (7.31)), we obtain

$$\begin{aligned} \| \prod_{j=1}^d \check{f}_j \|_{L^{q/d}(B_0(R))} &\lesssim R^{\frac{\alpha}{2}-\frac{d}{4}} \left[R^{-\frac{d}{2}} \int_{B_0(R)} \left(\prod_{j=1}^d \sum_{D_j} R^{-1/2} |(\tilde{f}_{j,D_j})^\vee|^2 * \mathbf{1}_{T_j}(x) \right)^{q/(2d)} dx \right]^{d/q} \\ &\lesssim R^{\alpha/2-d/2+\varepsilon/4} \prod_{j=1}^d \left(\sum_{D_j} \| \tilde{f}_{j,D_j} \|^2_{L^2(A_j^{\sqrt{R}})} \right)^{1/2} \\ &\lesssim R^{\alpha/2-d/2+\varepsilon/4} \prod_{j=1}^d \| f_j \|_{L^2(A_j^{\sqrt{R}})} = R^{\alpha/2-d/2+\varepsilon/4} \prod_{j=1}^d \| f_j \|_{L^2(A_j^R)} . \end{aligned}$$

In the penultimate inequality we used Kakeya and then Plancherel, and in the final inequality, the pointwise comparability $|\tilde{f}_{j,D_j}| \sim |f_{j,D_j}|$ and then the almost disjointness of the f_{j,D_j} to take the D_j-sum into the $L^2(A_j^{\sqrt{R}})$-norm. This concludes the proof of Proposition 7.34. □

8. Restriction estimates via reverse Littlewood–Paley inequalities and the Kakeya conjecture

Before we introduce the third tool commonly used to prove restriction estimates, let us discuss another possible approach to prove localized inequalities of the form (for $\mathbb{P}^{d-1}$ for the sake of concreteness)

$$\|(Fd\sigma)^\vee\|_{L^{2d/(d-1)}(B(x_0,R))} \lesssim R^\varepsilon \|F\|_{L^{2d/(d-1)}(\mathbb{P}^{d-1},d\sigma)}$$

for $R \gg 1$ and any $x_0 \in \mathbb{R}^d$. As we saw earlier, the above estimate can actually be reduced to

$$\|\check{G}\|_{L^{2d/(d-1)}(B(x_0,R))} \lesssim R^{\varepsilon-(d+1)/(2d)} \|G\|_{L^{2d/(d-1)}(\mathcal{N}_{1/R}(\mathbb{P}^{d-1}))}$$

for any $\check{G}$ with smooth Fourier support contained in $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$ (see Lemma 6.1). To make things simpler, let us only consider smooth functions $f \equiv \hat{G}$ with $\hat{f} = G$ belonging to the unit ball in $L^\infty(\mathcal{N}_{1/R}(\mathbb{P}^{d-1}))$, i.e., we are aiming to prove

$$\|f\|_{L^{2d/(d-1)}(B_R)} \lesssim R^{\varepsilon-1}\,.$$

Of course, this is a weaker statement (by Hölder's inequality), but by symmetry considerations one can actually show these statements are equivalent to each other.

We are now going to decompose $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$ into a collection of "slabs" $\theta \subseteq \mathbb{R}^d$, i.e., essentially disjoint curved regions with dimension $R^{-1/2} \times R^{-1/2} \times \cdots \times R^{-1}$. An explicit way to do this is to cover $[-1,1]^{d-1}$ with $2R^{-1/2} \times R^{-1/2} \times \cdots \times R^{-1/2}$ cubes $\{Q\}$ whose centers lie in the lattice $R^{-1/2}\mathbb{Z}^{d-1}$ and define each θ by

$$\theta = \{(\xi', \eta + |\xi'|^2) : \xi' \in Q_\theta, |\eta| \lesssim R^{-1}\}$$

for some choice of $Q_\theta \in \{Q\}$. We emphasize once more that it is important that the slabs are only essentially disjoint, i.e., they have some finite overlap which will also become manifest in a moment. In fact, the finite overlap allows us to construct a partition of unity of $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$ which is adapted to the family of slabs. Another consequence of this construction (and the curvature of $\mathbb{P}^{d-1}$) is the following observation concerning the set Ω of normals of these slabs.

Lemma 8.1. *The normals of the above slabs are $R^{-1/2}$-separated.*

Proof. For $j = 1, 2$, assume $(\xi_j', |\xi'|^2) \in \mathbb{P}^{d-1}$ and let $\nu_j = \nabla|\xi'|^2/\|\nabla|\xi'|^2\|$ (with $\nabla|\xi_j'|^2 = (2\xi_j', -1)$) denote the unit normal of $\mathbb{P}^{d-1}$ at $(\xi_j', |\xi'|^2)$. Then, by Cauchy–Schwarz and $|\xi_1' - \xi_2'| \sim R^{-1/2}$,

$$\nu_1 \cdot \nu_2 = \frac{4\xi_1' \cdot \xi_2' + 1}{(4|\xi_1'|^2+1)^{1/2}(4|\xi_2'|^2+1)^{1/2}} \leq \left(1 - \frac{4|\xi_1' - \xi_2'|^2}{(4|\xi_1'|^2+1)(4|\xi_2'|^2+1)}\right)^{1/2} \leq (1 - AR^{-1})^{1/2}\,.$$

Thus, we obtain

$$|\nu_1 - \nu_2|^2 = 2(1 - \nu_1 \cdot \nu_2) \gtrsim R^{-1}$$

by the mean value theorem. If $\nu_1 \cdot \nu_2 < 0$, the above difference is even $\mathcal{O}(1)$. $\square$

We will now decompose f using the partition of unity that is given by the slabs θ, i.e.,

$$f = \sum_{\theta: R^{-1/2}-\text{slab}} f_\theta \quad \text{where } \hat{f}_\theta = \hat{f}\mathbf{1}_\theta\,.$$

Our goal is then to prove

$$\left\| \sum_{\theta: R^{-1/2}-\text{slab}} f_\theta \right\|_{L^{2d/(d-1)}(B(x_0,R))} \lesssim R^{\varepsilon-1}\,.$$

In fact, we will show the ostensibly stronger estimate

$$\Big\|\sum_{\theta:R^{-1/2}-\text{slab}} f_\theta\Big\|_{2d/(d-1)} \lesssim R^{\varepsilon-1}\,.$$

The main difficulty is to understand the cancellation properties between the individual f_θ. Therefore, our first goal is to replace the ℓ^1 quantity $\sum_\theta |f_\theta|$ by the ℓ^2 quantity $\left(\sum_\theta |f_\theta|^2\right)^{1/2}$ which has the effect of separating the contributions from individual f_θ whilst accounting for any destructive interference. Unfortunately, such a strong relationship has not been obtained yet, which is why we only have the following

Conjecture 8.2 (Reverse Littlewood–Paley inequality for slabs)**.** *Suppose f has frequency support in $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$. Then*

$$\|f\|_{L^p(\mathbb{R}^d)} \lesssim R^\varepsilon \Bigg\|\Bigg(\sum_{\theta:R^{-1/2}-\text{slab}} |f_\theta|^2\Bigg)^{1/2}\Bigg\|_{L^p(\mathbb{R}^d)} \qquad \text{for } 2\le p\le \frac{2d}{d-1}\,. \tag{8.1}$$

For even Hölder exponents, the reverse square function estimate can be proved under the condition that Minkowski sums of sets have only bounded overlap.

Definition 8.3. Let $(\Omega_j)_{j=1}^n$ be a sequence of sets in $\mathbb{R}^d$. We say that "ξ lies in at most $A(\xi)\in\mathbb{N}$ of the Ω_j" whenever the maximal number of Ω_j which contain ξ is given by $A(\xi)$, i.e.,

$$A(\xi) := \sup\{\text{number of } \Omega_j' s \text{ containing } \xi\}\,.$$

Then we have the following

Proposition 8.4 (Reverse L^2 and L^4 square function estimates)**.** *Let $f_1,...,f_n\in\mathcal{S}(\mathbb{R}^d)$ have Fourier support in sets $\Omega_1,...,\Omega_n\subseteq\mathbb{R}^d$, respectively. Then we have the following assertions.*

1. *(Almost orthogonality) If the sets $\Omega_1,...,\Omega_n$ have overlap at most A_2, (i.e., every ξ lies in at most $A_2\in\mathbb{N}$ of the Ω_j, i.e., $\sup_{\xi\in\mathbb{R}^d} A(\xi)\le A_2$) for some $A_2>0$, then*

$$\|\sum_{j=1}^n f_j\|_{L^2(\mathbb{R}^d)} \le A_2^{1/2}\|(\sum_{j=1}^n |f_j|^2)^{1/2}\|_{L^2(\mathbb{R}^d)}\,.$$

2. *(Almost bi-orthogonality) If the (n^2) sum sets $\Omega_i+\Omega_j := \{\xi+\xi' :\ \xi\in\Omega\,,\xi'\in\Omega'\}$ with $i,j\in\{1,...,n\}$ have overlap at most A_4 for some $A_4>0$, then*

$$\|\sum_{j=1}^n f_j\|_{L^4(\mathbb{R}^d)} \le A_4^{1/4}\|(\sum_{j=1}^n |f_j|^2)^{1/2}\|_{L^4(\mathbb{R}^d)}\,.$$

Remarks 8.5. (1) Clearly, the above theme can be generalized for L^{2p} with $p\in\mathbb{N}$, if one assumes that the sum sets $\sum_{j=1}^p \Omega_j$ have overlap at most A_{2p}, see, e.g., Gressman–Guo–Pierce–Roos–Yung [98].

(2) By using $f_i\overline{f_j}$ in place of f_if_j, one can also establish a variant of (2) in Proposition 8.4 where the sum set $\Omega_i+\Omega_j$ is replaced by the difference set $\Omega_i-\Omega_j := \{\xi-\xi' :\ \xi\in\Omega\,,\xi'\in\Omega'\}$.

Proof. (1) For $p=2$ this is an immediate consequence of Plancherel's theorem, pointwise Cauchy–Schwarz

$$(\sum_{j=1}^n |\hat f_j(\xi)|)^2 \le A_2(\sum_{j=1}^n |\hat f_j(\xi)|^2)$$

(since the sets are only finitely overlapping) and Fubini. More precisely,

$$\|\sum_{j=1}^{n} f_j\|_2 = \|\sum_{j=1}^{n} \hat{f}_j\|_2 \leq A_2^{1/2} \|(\sum_{j=1}^{n} |\hat{f}_j|^2)^{1/2}\|_2 = A_2^{1/2} \|(\sum_{j=1}^{n} |f_j|^2)^{1/2}\|_2 \,.$$

(2) Writing

$$\|\sum_{j=1}^{n} f_j\|_4^2 = \|\sum_{i,j=1}^{n} f_i f_j\|_2$$

and

$$\|(\sum_{i,j=1}^{n} |f_i f_j|^2)^{1/2}\|_2 = \|(\sum_{j=1}^{n} |f_j|^2)^{1/2}\|_4^2$$

it becomes obvious that this assertion follows from what we have just shown. More precisely, using the fact that $f_i f_j$ has Fourier support in the Minkowski sum $\Omega_i + \Omega_j$ (by the convolution theorem, i.e., supp $\hat{f}_1 * \hat{f}_2 \subseteq$ supp $\hat{f}_1 +$ supp $\hat{f}_2$) and the fact that these sums only overlap finitely, we can apply Cauchy–Schwarz in the i, j-summation, i.e.,

$$(\sum_{i,j=1}^{n} \widehat{f_i f_j}(\xi))^2 \leq A_4 \sum_{i,j=1}^{n} |\widehat{f_i f_j}|^2(\xi)$$

and conclude as before using Plancherel. □

In Appendix B we will apply the above observation to review a classical argument due to Córdoba which proves the reverse Littlewood–Paley inequality (and thereby the restriction conjecture) in $d = 2$ when $p = 4$. (Recall that we already presented in Subsection 7.4 an alternative proof of two-dimensional restriction relying on bilinear techniques.)

Remark 8.6. In fact, an argument of Carbery [38] shows that the hypothesized square function estimate (8.1) implies the Kakeya conjecture and, consequently, the restriction conjecture. Attempting to prove the whole restriction conjecture from this direction seems a quite optimistic strategy as (8.1) appears to be very powerful and in all likelihood considerably more difficult than the restriction conjecture.

From now on, we will assume that the reversed Littlewood–Paley inequality holds. The frequency localization onto the slabs leads (by the uncertainty principle) to a localization to dual tubes which is called *wave packet decomposition* and which will be discussed in the next section. Let us anyway anticipate already the main result of that section, Lemma 9.2, which says that there exist constants f_T and a collection $\mathbb{T}(\theta)$ of tubes dual to the slab θ (which is centered at $\xi_\theta \in \mathbb{R}^d$) which cover $\mathbb{R}^d$ such that

$$f_\theta(x) = \sum_{T \in \mathbb{T}(\theta)} f_T \psi_T(x) \,,$$

where $\psi_T(x) = |T|^{-1} \mathrm{e}^{-2\pi i x \cdot \xi_\theta} \varphi_T(x)$ is a so-called wave packet associated to T. Here, $\varphi_T = \varphi \circ a_T^{-1}$ where φ is a Schwartzian bump function centered at the origin with supp $\hat{\varphi} \subseteq [-1/2, 1/2]^d$ and a_T is an affine transformation whose linear part has determinant $|T|$ and maps $[-1/4, 1/4]^d$ bijectively to T. (Recall that for an invertible linear map $S : \mathbb{R}^d \to \mathbb{R}^d$, one has

$$\widehat{f \circ S_T} = |\det(S_T)|^{-1} \hat{f} \circ S^{-t}$$

where $S{-}t$ denotes the inverse transpose of S.)

Now, applying the Littlewood–Paley conjecture together with the wave packet decomposition implies that it suffices to bound (noting $|T| \sim R^{(d+1)/2}$)

$$R^{-(d+1)/2} \left\| \left(\sum_{\theta: R^{-1/2}-\text{slab}} \left[\sum_{T \in \mathbb{T}(\theta)} |f_T||\varphi_T| \right]^2 \right)^{1/2} \right\|_{L^{2d/(d-1)}(\mathbb{R}^d)} .$$

We will do so by replacing the "smooth indicator function" φ_T (which decays rapidly away from T) by a sharp cut-off $\mathbf{1}_T$ and afterwards applying the Kakeya conjecture 15.1. In this context (using $|T| \sim R^{(d+1)/2}$ and that the number of $R^{-1/2}$-separated slabs covering $\mathbb{P}^{d-1}$ is $\mathcal{O}(R^{(d-1)/2})$), the conjecture says

$$\left\| \sum_{\theta: R^{-1/2}-\text{slab}} \mathbf{1}_{T_\theta} \right\|_{L^{d/(d-1)}(\mathbb{R}^d)} \lesssim R^{\varepsilon+d-1} .$$

We begin with the replacement of φ_T by $\mathbf{1}_T$. For this, let $\ell \in \mathbb{Z}^d$ and $\mathbf{1}_{T,\ell}$ denote the characteristic function of the rectangle $a_T\left([-1/4,1/4]^d + \ell/2\right)$. Thus, the $\mathbf{1}_{T,\ell}$ yield a rough partition of unity of $\mathbb{R}^d$.

Lemma 8.7. *With the above notation, the estimate*

$$\left\| \left(\sum_{\theta: R^{-1/2}-\text{slab}} \left[\sum_{T \in \mathbb{T}(\theta)} |f_T||\varphi_T| \right]^2 \right)^{1/2} \right\|_{L^{2d/(d-1)}(\mathbb{R}^d)}$$
$$\lesssim \sum_{\ell \in \mathbb{Z}^d} (1+|\ell|)^{-(d+1)} \left\| \left(\sum_{\theta: R^{-1/2}-\text{slab}} \left[\sum_{T \in \mathbb{T}(\theta)} |f_T| \mathbf{1}_{T,\ell} \right]^2 \right)^{1/2} \right\|_{L^{2d/(d-1)}(\mathbb{R}^d)}$$

holds.

Proof. This follows from the rapid decay of φ, i.e.,

$$|\varphi_T(x)| = \sum_{\ell \in \mathbb{Z}^d} \varphi_T(x) \mathbf{1}_{T,\ell}(x) \lesssim \sum_{\ell \in \mathbb{Z}^d} \frac{\mathbf{1}_{T,\ell}(x)}{\left(1+|a_T^{-1}(x)|\right)^{d+1}} \lesssim \sum_{\ell \in \mathbb{Z}^d} \frac{\mathbf{1}_{T,\ell}(x)}{(1+|\ell|)^{d+1}}$$

and a two-fold application of Minkowski's inequality (first in the ℓ^2-norm and afterwards in the $L^{2d/(d-1)}$-norm). □

Since the supports of the $\mathbf{1}_{T,\ell}$ are essentially disjoint as T varies over $\mathbb{T}(\theta)$ (i.e., $\mathbf{1}_{T_1,\ell}(x)\mathbf{1}_{T_2,\ell}(x) = 0$ for almost all $T_1, T_2 \in \mathbb{T}(\theta)$), one has

$$\left[\sum_{T \in \mathbb{T}(\theta)} |f_T| \mathbf{1}_{T,\ell} \right]^2 \lesssim \sum_{T \in \mathbb{T}(\theta)} |f_T|^2 \mathbf{1}_{T,\ell} .$$

That means that the $L^{2d/(d-1)}$-norm (of the right side appearing in the inequality of the above lemma) for a fixed $\ell \in \mathbb{Z}^d$ can be bounded by

$$\left\| \sum_{\theta: R^{-1/2}-\text{slab}} \sum_{T \in \mathbb{T}(\theta)} |f_T|^2 \mathbf{1}_{T,\ell} \right\|_{L^{d/(d-1)}(\mathbb{R}^d)}$$

This means that it suffices to show that this expression is $\mathcal{O}(R^{(d-1)/2})$. Using the information on $\sum_T |f_T|^2$ from the wave packet decomposition and our initial hypothesis that $\hat{f}$ belongs to the unit ball of $L^\infty(\mathcal{N}_{1/R}(\mathbb{P}^{d-1}))$, we have

$$\sum_{T\in\mathbb{T}(\theta)} |f_T|^2 \lesssim 1\,,$$

i.e., there exists a sequence $(c_T)_{T\in\mathbb{T}(\theta)}$ of non-negative real numbers such that

$$\sum_{T\in\mathbb{T}(\theta)} c_T^2 \lesssim 1$$

and that

$$\left\|\sum_{\theta:R^{-1/2}-\text{slab}}\sum_{T\in\mathbb{T}(\theta)} |f_T|^2\mathbf{1}_{T,\ell}\right\|_{L^{d/(d-1)}(\mathbb{R}^d)} \lesssim \left\|\sum_{\theta:R^{-1/2}-\text{slab}}\sum_{T\in\mathbb{T}(\theta)} c_T\mathbf{1}_{T,\ell}\right\|_{L^{d/(d-1)}(\mathbb{R}^d)}.$$

Lemma 8.8. *With the above notation, we have*

$$\left\|\sum_{\theta:R^{-1/2}-\text{slab}}\sum_{T\in\mathbb{T}(\theta)} c_T\mathbf{1}_{T,\ell}\right\|_{L^{d/(d-1)}(\mathbb{R}^d)} \lesssim \mathbb{E}\left\|\sum_{\theta:R^{-1/2}-\text{slab}} \mathbf{1}_{T_\theta,\ell}\right\|_{L^{d/(d-1)}(\mathbb{R}^d)}$$

for any $\ell\in\mathbb{Z}^d$ and choice $(T_\theta)\in\prod_\theta\mathbb{T}(\theta)$ [12].

Believing this lemma for the moment, the argument is concluded by applying the hypothesized Kakeya estimate

$$\left\|\sum_{\theta:R^{-1/2}-\text{slab}} \mathbf{1}_{T_\theta,\ell}\right\|_{L^{d/(d-1)}(\mathbb{R}^d)} \lesssim R^{\varepsilon+d-1} \tag{8.2}$$

which is valid for every choice $\ell\in\mathbb{Z}^d$ and $(T_\theta)\in\prod_\theta\mathbb{T}(\theta)$. We conclude the section with the

Proof of Lemma 8.8. (1) Consider randomly selecting a sequence of rectangles, one for each direction θ. Each T is chosen from $\mathbb{T}(\theta)$ with a probability c_T [13]. This means that we constructed a probability space $\prod_\theta\mathbb{T}(\theta)$ where a sequence of rectangles (i.e., a singleton $\{(T_\theta)\}$) is picked with the probability $\prod_\theta c_{T_\theta}$.

(2) For a fixed $x\in\mathbb{R}^d$, consider the random variable $\sum_\theta \mathbf{1}_{T_\theta,\ell}(x)$ which counts the number of rectangles of the above randomly picked sequence (T_θ) for which $x\in\operatorname{supp}\mathbf{1}_{T_\theta,\ell}$. The expectation value (with respect to the "probability space" $\mathbb{T}(\theta)$) that $x\in\operatorname{supp}\mathbf{1}_{T_\theta,\ell}$ holds for a given θ is given by

$$\mathbb{E}\mathbf{1}_{T_\theta,\ell}(x) = \sum_{T\in\mathbb{T}(\theta)} c_T\mathbf{1}_{T_\theta,\ell}(x)\,.$$

Thus, by the linearity of the expectation, one has

$$\mathbb{E}\left[\sum_{\theta:R^{-1/2}-\text{slab}} \mathbf{1}_{T_\theta,\ell}(x)\right] = \sum_{\theta:R^{-1/2}-\text{slab}}\sum_{T\in\mathbb{T}(\theta)} c_T\mathbf{1}_{T_\theta,\ell}(x)\,.$$

[12] Here, (T_θ) is understood as a randomly picked sequence of rectangles, one for each direction θ. The space $\prod_\theta\mathbb{T}(\theta)$ is thus endowed with a probability measure which assigns the probability $\prod_\theta c_{T_\theta}$ to each singleton $\{(T_\theta)\}$.

[13] More precisely, consider a sequence of slabs $(\theta_j)_{j\in\mathbb{N}}$. Then for each slab θ_j, there is a sequence of rectangles $(T^n_{\theta_j})_{n\in\mathbb{N}}\in\mathbb{T}(\theta_j)$ (which covers $\mathbb{R}^d$) and the above c_T actually means $c_{T^n_{\theta_j}}$

Taking the $L^{d/(d-1)}$-norm of these expressions, we infer from Minkowski's inequality

$$\Big\| \sum_{\theta : R^{-1/2}-\text{slab}} \sum_{T \in \mathbb{T}(\theta)} c_T \mathbf{1}_{T_\theta,\ell}(x) \Big\|_{L^{d/(d-1)}(\mathbb{R}^d)} \leq \mathbb{E} \Big\| \sum_{\theta : R^{-1/2}-\text{slab}} \mathbf{1}_{T_\theta,\ell}(x) \Big\|_{L^{d/(d-1)}(\mathbb{R}^d)}$$

for any $\ell \in \mathbb{Z}^d$ and choice $(T_\theta) \in \prod_\theta \mathbb{T}(\theta)$. □

9. The wave packet decomposition

We will now present the third method which has been used over the last say ten years to obtain restriction estimates.

For the sake of illustration, assume that we wish to prove restriction estimates for the paraboloid $\mathbb{P}^{d-1}$ using restriction estimates via reverse Littlewood–Paley inequalities.

Recall that we covered $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$ with $R^{-1/2} \times \cdots \times R^{-1/2} \times R^{-1}$-slabs θ whose normal directions were $R^{-1/2}$-separated. This lead us to the decomposition $f = \sum_\theta f_\theta$ in Fourier space where $\hat{f}_\theta = \hat{f}\mathbf{1}_\theta$. By the uncertainty principle, localizing to a slab θ which is oriented in direction ω is equivalent to localizing to a dual tube T (or rather to a collection of such tubes covering $\mathbb{R}^d$) of dimensions $R^{1/2} \times \cdots \times R^{1/2} \times R$ in physical space which is oriented along ω as well. The functions which are going to localize to these tubes are called *wave packets* and can be thought of as smoothed out copies of Knapp examples. The goal of this section is to make the above intuition more precise.

Let T be some rectangle and a_T be an affine transformation whose linear part has determinant $|T|$ and which maps $[-1/4, 1/4]^d$ bijectively to T.

Let further φ be a Schwartzian bump function at the origin such that supp $\hat{\varphi} \subseteq [-1/2, 1/2]^d$ and $\hat{\varphi}|_{[-1/4,1/4]^d} = 1$. Define then $\varphi_T := \varphi \circ a_T^{-1}$ as the bump function on the tube T. (Recall that for an invertible linear map $S : \mathbb{R}^d \to \mathbb{R}^d$, one has

$$\widehat{f \circ S} = |\det(S)|^{-1} \hat{f} \circ S^{-t}$$

where S^{-t} denotes the inverse transpose of S.)

Finally, for a given slab θ, we denote by $\mathbb{T}(\theta)$ the finitely overlapping collection of tubes which are dual to θ, oriented along the direction of θ.

With the above notation, we are finally in position to define wave packets.

Definition 9.1 (Wave packets)**.** Let θ be an $R^{-1/2}$-slab centered at $\xi_\theta \in \mathbb{R}^d$. Let T, a_T, φ_T and $\mathbb{T}(\theta)$ as above. Then a wave packet associated to $T \in \mathbb{T}(\theta)$ is defined as

$$\psi_T(x) := |T|^{-1} e^{2\pi i x \cdot \xi_\theta} \varphi_T(x) \, .$$

Before we make the heuristics of the beginning of the section precise, the following crucial observations are in order.

(1) If a_T^* denotes the adjoint of the linear part of the affine transformation a_T, then $|\hat{\psi}_T(\xi)| \sim |\hat{\varphi}(a_T^*(\xi - \xi_\theta))$ and $\hat{\psi}_T$ is supported on a dilute of θ with $|\hat{\psi}_T|_\theta| = 1$.

(2) We have the support property $\{\xi \in \mathbb{R}^d : |\hat{\psi}_T(\xi)| = 1\} \subseteq (a_T^*)^{-1}\big([-1/4, 1/4]^d\big) + \xi_\theta$ where $(a_T^*)^{-1}\big([-1/4, 1/4]^d\big)$ is a rectangle dual to T.

Lemma 9.2 (Wave packet decomposition)**.** *Let $f \in C^\infty(\mathbb{R}^d)$ with Fourier support in $\mathcal{N}_{1/R}(\mathbb{P}^{d-1})$. Then for any $R^{-1/2}$-slab θ there exists a decomposition*

$$f_\theta(x) = \sum_{T \in \mathbb{T}(\theta)} f_T \psi_T(x)$$

where the constants f_T satisfy

$$\left(\sum_{T\in\mathbb{T}(\theta)} |f_T|^2\right)^{1/2} \lesssim \|\hat{f}_\theta\|_{L^2_{\mathrm{avg}}(\theta)} \,.$$

Here, the averaged L^p norm $\|\cdot\|_{L^p_{\mathrm{avg}}(\Omega)}$ for some subset $\Omega \subseteq \mathbb{R}^d$ of finite Lebesgue measure is defined as

$$\|f\|_{L^p_{\mathrm{avg}}(\Omega)} := |\Omega|^{-1/p}\|f\|_{L^p(\Omega)} \,. \tag{9.1}$$

Proof. Denote by T_0 the $R^{1/2} \times \cdots \times R^{1/2} \times R$-rectangle oriented along θ and centered at 0. Then

$$g_\theta(\xi) := \hat{f}_\theta((a^*_{T_0})^{-1}\xi + \xi_\theta)$$

is supported on $[-1/2, 1/2]^d$ and can be thought of as a function on the torus $\mathbb{T}^d = [-1/2, 1/2]^d$. That means, that it can be expanded in a Fourier series whose Fourier coefficients u_k satisfy

$$\sum_{k\in\mathbb{Z}^d} |u_k|^2 = \|g_\theta\|^2_{L^2([-1/2,1/2]^d)} \lesssim \|\hat{f}_\theta\|^2_{L^2_{\mathrm{avg}}(\theta)} \,.$$

Therefore,

$$\hat{f}_\theta(\xi) = g_\theta(a^*_{T_0}(\xi - \xi_\theta)) = \sum_{k\in\mathbb{Z}^d} u_k \mathrm{e}^{-2\pi i k\cdot a^*_{T_0}(\xi-\xi_\theta)} \quad \text{for } \xi \in (a^*_{T_0})^{-1}\left([-1/2, 1/2]^d\right) + \xi_\theta \,.$$

On the other hand, we saw in our earlier considerations that the function $\hat{\varphi}(a^*_{T_0}(\xi - \xi_\theta))$ equals one on supp $\hat{f}_\theta$ and is itself supported on $(a^*_{T_0})^{-1}\left([-1/2, 1/2]^d\right) + \xi_\theta$. But that means that the last equality can also be written as

$$\hat{f}_\theta(\xi) = g_\theta(a^*_{T_0}(\xi - \xi_\theta)) = \sum_{k\in\mathbb{Z}^d} u_k \mathrm{e}^{-2\pi i k\cdot a^*_{T_0}(\xi-\xi_\theta)}\hat{\varphi}(a^*_{T_0}(\xi - \xi_\theta)) \quad \text{for } \xi \in \mathbb{R}^d \,.$$

Performing an inverse Fourier transform on the last expression then leads to

$$f_\theta(x) = \sum_k u_k |\det a^{-1}_{T_0}|\mathrm{e}^{2\pi i x\cdot\xi_\theta}\varphi_{T_0}(x - a_{T_0}k) = \mathrm{const}\sum_k u_k \psi_{T_0+a_{T_0}k}(x) \,.$$

The proof is concluded by noting that $\mathbb{T}(\theta)$ is just the collection of all rectangles of the form $\{T_0 + a_{T_0}k\}_{k\in\mathbb{Z}^d}$. □

10. The polynomial method

We follow Demeter [63, Ch. 8] and refer to Guth's extensive treatment [102].

One notable feature of affine subspaces, such as points and lines, is their lack of scale. Can polynomials be used to count cubes and tubes? How about to estimate integrals involving complicated expressions? Dvir's proof [69] of the finite field Kakeya conjecture introduced a robust way of counting structures (e.g., special points, lines) with the aid of polynomials. Another satisfactory answer came in the form of the resolution of the endpoint multilinear Kakeya conjecture by Guth [100].

Perhaps even more surprising is the fact that Guth also managed to tailor the polynomial method to produce significant progress on the restriction conjecture, a highly oscillatory problem that is not a priori formulated as a counting problem. His approach can be summarized as follows. Use a polynomial P of small degree to partition the spatial ball $B(0, R)$ into cells where $|Ef|^p$ has roughly the same mass. On $B(0, R)$, the function Ef is the sum of wave packets whose mass is concentrated inside tubes. The key is to understand how the various wave packets interact with the cells. The contribution from the tubes that lie roughly tangent to the zero set $Z(P)$ of the polynomial P is estimated directly, using both counting arguments for tubes and the oscillatory properties of Ef. The contribution from the other tubes is estimated using a well-crafted induction hypothesis, exploiting only the algebraic properties of the cells and the L^2 orthogonality of the wave packet decomposition.

Here, we will merely introduce the polynomial partitioning method and refer to Demeter [63] and Guth [102] for applications.

10.1. **Polynomial partitioning.** The main idea of polynomial partitioning is to *simultaneously bisect N masses in $\mathbb{R}^n$ using a polynomial of low degree.* The low degree of the polynomial allows one to exploit the fact that any line can interact the zero set $Z(P)$ of a polynomial only a few times.

The first building block is the following classical theorem from algebraic topology.

Theorem 10.1 (Borsuk–Ulam). *Let $f : \mathbb{S}^n \to \mathbb{R}^n$ be continuous. Then the following two statements hold and are equivalent to each other.*

(1) There is $x_0 \in \mathbb{S}^n$ such that $f(-x_0) = f(x_0)$.
(2) If $f(-x) = -f(x)$ for all $x \in \mathbb{S}^n$, then there is $x_0 \in \mathbb{S}^n$ such that $f(x_0) = 0$.

Remarks 10.2. (1) The case $n = 1$ can be illustrated by saying that there always exist a pair of opposite points on the Earth's equator with the same temperature. The same is true for any circle. This assumes the temperature varies continuously in space.

(2) The case $n = 2$ is often illustrated by saying that at any moment, there is always a pair of antipodal points on the Earth's surface with equal temperatures and equal barometric pressures, assuming that both parameters vary continuously in space.

Theorem 10.1 is the key to prove the following "ham sandwich theorem" due to Stone and Tukey [170].

Theorem 10.3. *Let $\mathcal{F}$ be a real vector space with dimension* $\dim(\mathcal{F})$ *consisting of real-valued, continuous functions on $\mathbb{R}^n$. (For instance the space of polynomials with given degree.) Let $f_0 \equiv 0$ denote the trivial function. For $N < \dim(\mathcal{F})$ let $\mu_1, ..., \mu_N$ be finite Borel measures on $\mathbb{R}^n$ such that for each $f \in \mathcal{F} \setminus \{f_0\}$, one has $\mu_j(Z(f)) = 0$ for all $j = 1, ..., N$.*

Then there is $f \in \mathcal{F} \setminus \{f_0\}$ such that for each $j = 1, ..., N$ one has

$$\mu_j(\{x \in \mathbb{R}^n : f(x) > 0\}) = \mu_j(\{x \in \mathbb{R}^n : f(x) < 0\}) . \tag{10.1}$$

We now apply Theorem 10.3 to the space $\mathcal{F} = \mathrm{Poly}_D(\mathbb{R}^n)$ of polynomials on $\mathbb{R}^n$ of degree $\leq D$. In the following estimates, the dependence on D will be important, while that on the ambient space dimension is swept under the rug as always. We start with two preliminary observations.

(1) A simple counting argument reveals that this space has a fairly large dimension, namely

$$\dim(\mathrm{Poly}_D(\mathbb{R}^n)) = \binom{D+n}{n} \sim_n D^n \tag{10.2}$$

(2) The zero set of any polynomial P has zero Lebesgue measure (as can be seen from a simple induction argument).

Corollary 10.4 (Polynomial ham sandwich). *Let $\mu_1, ..., \mu_N$ be finite Borel measures on $\mathbb{R}^n$ that are absolutely continuous with respect to the Lebesgue measure. Then there is a non-trivial polynomial $P \in \mathrm{Poly}_D(\mathbb{R}^n)$ with $D \lesssim_n N^{1/n}$ such that for each $j = 1, ..., N$ we have*

$$\mu_j(\{x \in \mathbb{R}^n : P(x) > 0\}) = \mu_j(\{x \in \mathbb{R}^n : P(x) < 0\}) . \tag{10.3}$$

Iterating this corollary $\sim D^n$ times allows us to split any "mass" into N equal pieces.

Theorem 10.5 (Guth [103])**.** *Let $0 \leq W \in L^1(\mathbb{R}^n)$. Then for each degree D there is $0 \neq P \in \mathrm{Poly}_D(\mathbb{R}^n)$ such that $\mathbb{R}^n \setminus Z(P)$ is the disjoint union of $\sim_n D^n$ open sets $\{O_j\}_j$ (called "cells") such that all integrals $\int_{O_j} W$ are equal. Moreover, each line in $\mathbb{R}^n$ can intersect at most $D+1$ cells O_j.*

Example 10.6. Let $W = \mathbf{1}_{[0,1]^n}$. Then the polynomial

$$P(x_1, ..., x_n) = \prod_{j=1}^{n} \prod_{m=1}^{M-1} \left(x_j - \frac{m}{M}\right) \tag{10.4}$$

has degree $D = n(M-1)$. Moreover, $\mathbb{R}^n \setminus Z(P)$ has M^n connected components O_j satisfying $\int_{O_j} W = M^{-n}$ for all $j = 1, ..., M^n$.

For technical reasons, it is helpful in our arguments later to use non-singular polynomials.

Definition 10.7 (Non-singular points and polynomials)**.** A point $x \in \mathbb{R}^n$ is called non-singular for a function $P : \mathbb{R}^n \to \mathbb{C}$, if $\nabla P \neq 0$. A polynomial P is called non-singular, if every point in $Z(P)$ is non-singular for P.

Remarks 10.8. (1) The neighborhood of $Z(P)$ close enough to a non-singular point is a smooth hypersurface.

(2) In most applications, it suffices for the integrals W to be comparable, rather than equal to each other. Using a density argument, this can be achieved by using only nonsingular polynomials P_i in the proof of Theorem 10.5 (cf. [103]). A product of nonsingular polynomials has the property that its non-singular points are dense in its zero set. This additional regularity makes some of the arguments easier. We make this remark precise in the following theorem.

Theorem 10.9. *Suppose $0 \leq W \in L^1(\mathbb{R}^n)$. Then for any D there exists a non-zero polynomial P of degree $\leq D$ so $R^n \setminus Z(P)$ is a disjoint union of $\sim_n D^n$ open sets O_j. Moreover, the integrals $\int_{O_j} W$ agree up to a factor of 2. Finally, the polynomial P is a product of non-singular polynomials.*

11. Induction on scales

12. Adapting Wolff's argument to the paraboloid

13. Connection to PDEs

13.1. **Original Strichartz estimates.** Strichartz [172, §3] observed that restriction theorems immediately yield estimates on the L^p norms of solutions to certain dispersive PDEs, in particular the free Schrödinger equation, the Klein–Gordon equation, and the acoustic wave equation. We begin this section by giving classic bounds on $\|u\|_{L_x^p}$ for the free Schrödinger equation. We will then generalize these estimates to mixed norm estimates which are invaluable to prove global well-posedness of nonlinear dispersive equations such as the cubic nonlinear Schrödinger equation.

The main theorem of this section is the following result [172, Corollary 1]. The full range of Strichartz estimates were proven by Keel and Tao [119], whereas non-endpoint results were obtained by Ginibre and Velo [95] and Yajima [203].

Theorem 13.1 (Strichartz estimate for the free Schrödinger equation)**.** *Let $u(x,t)$ be the solution of the inhomogeneous, free Schrödinger equation*

$$\begin{aligned} i\frac{\partial u}{\partial t}(x,t) + \lambda\Delta_x u(x,t) &= g(x,t) \\ u(x,0) &= f(x) \end{aligned} \tag{13.1}$$

for $x \in \mathbb{R}^d$*,* $t \in \mathbb{R}$*, and* $\lambda \in \mathbb{R}\backslash\{0\}$*. Assume* $f \in L^2(\mathbb{R}^d)$ *and* $g \in L^p(\mathbb{R}^{d+1})$ *for* $p = 2(d+2)/(d+4)$*. Then* $u \in L^q(\mathbb{R}^{d+1})$ *for* $q = 2(d+2)/d$ *and* $\|u\|_q \leq a(\|f\|_2 + \|g\|_p)$*.*

Proof. It is well known that (13.1) has a unique solution which can be written as

$$u(x,t) = \int_0^t \mathrm{e}^{i\lambda(t-s)\Delta} g(x,s)\, ds + a \int_{\mathbb{R}^d} \hat{f}(\xi)\mathrm{e}^{-i(x\cdot\xi+\lambda\xi^2\cdot 2t)}\, d\xi$$

by performing a Fourier transform and applying Duhamel's formula. The estimate for the second term is then an immediate consequence of the restriction theorem

$$\|(F d\sigma)^\vee\|_{L^q(\mathbb{R}^{d+1})} \lesssim \|F\|_{L^2(S)}$$

where the manifold S is the paraboloid starting at the origin, i.e.,

$$S = \{(x,t) \in \mathbb{R}^{d+1} : R(x,t) := t - \lambda|x|^2 = 0\}\,.$$

To estimate the first term we use $\|\mathrm{e}^{it\Delta}\|_{2\to 2} = 1$ (by unitarity) and $\|\mathrm{e}^{it\Delta}\|_{1\to\infty} \lesssim |t|^{-d/2}$ (by the fundamental solution of the free Schrödinger equation). Thus, by interpolation,

$$\|\mathrm{e}^{it\Delta}\|_{p\to q} \lesssim |t|^{-d\left(\frac{1}{p}-\frac{1}{2}\right)} = |t|^{-d/(d+2)}\,.$$

Thus, with $r = d/(d+2)$ (i.e., $1/p - 1/q = 1 - r = 2/(d+2)$), and the Hardy–Littlewood–Sobolev inequality,

$$\left\|\int_0^t \mathrm{e}^{i\lambda(t-s)\Delta} g(\cdot,s)\, ds\right\|_{L^q(\mathbb{R}^d)} \lesssim \int_0^t |t-s|^{-r}\|g(\cdot,s)\|_{L^p(\mathbb{R}^d)}\, ds \lesssim \|g\|_{L^p(\mathbb{R}^{d+1})}\,,$$

which was asserted. □

Remark 13.2. One should compare the last inequality with the Christ–Kiselev lemma [47] which says the following.

Let X, Y be Banach spaces, I be a time interval, and $K \in C^0(I \times I : B(X \to Y))$ be a kernel taking values in the space of bounded operators from X to Y. Suppose $1 \leq p < q \leq \infty$ is such that

$$\left\|\int_I K(t,s) f(s)\, ds\right\|_{L_t^q(I:Y)} \lesssim \|f\|_{L_t^p(I:X)}$$

for all $f \in L^p(I : X)$. Then, for any $s < t$, one also has

$$\left\|\int_{s\in I: s<t} K(t,\tau) f(\tau)\, d\tau\right\|_{L_t^q(I:Y)} \lesssim \|f\|_{L_t^p(I:X)}$$

The principle that motivates this lemma is that if an operator is known to be bounded from one space to another, then any "reasonable localization" (in this case to the causal region $s < t$ of time interactions) of that operator should be bounded as well. Unfortunately, the condition $p < q$ is necessary.

The proof of this lemma as it was formulated here can be found in Smith and Sogge [158, Lemma 3.1] or Tao [182]. A slight variation thereof was used by Mizutani and Yao [139, Appendix C].

For Strichartz estimate for the Schrödinger equation with scalar potentials, we refer to the recent paper by Kim–Seo–Seok [122] and the vast list of references therein. Let us in particular emphasize the groundbreaking works by Bouclet and Mizutani [14] and Burq et al [32, 33] for Schrödinger operators with critical singularities and critical decay (in particular of Hardy's type). The generalization of Strichartz inequalities to fractional Hardy operators was carried out by Mizutani and Yao [139, Theorems 1.7 and 1.8].

13.2. **Global well-posedness of the cubic NLS in $d = 2$.**

13.2.1. *Non-linear dispersive equations.* Let us discuss some immediate consequences of the restriction conjecture regarding evolution equations. Examples for such equations are the heat equation

$$\partial_t u - \Delta u = 0\,,$$

the wave equation

$$\partial_t^2 u - \Delta u = 0\,,$$

and Schrödinger's equation

$$i\partial_t u - \Delta u = 0\,.$$

There are many other important evolution equations such as the Euler or the Navier–Stokes equation which describe the motion of fluids.

For evolution equations, the natural problem to study is the *Cauchy problem* (as opposed to, say, the Dirichlet problem). We specify initial data $u(0) = f$ (and, for the wave equation also the initial velocity $\partial_t u(0) = g$) and ask for the solution u at a later time t. There are three fundamental questions that one can ask about such equations.

- Existence: does a solution $u(t)$ exist at all? In what sense (weak, strong, classical) is it a solution? Does it exist for all times, or just for a finite time interval?
- Uniqueness: can there be more than one solution with the same initial data? Are there some extra conditions (e.g., regularity conditions) one needs to impose to force uniqueness? If there are still several solutions, is there a "good", or "physically relevant" solution that is somehow "better" than the others?
- Stability: suppose we perturb the initial data slightly. How does this affect the solution? More precisely, does the solution *depend continuously* on the data (as measured in some Banach space norm, for instance)?

An equation is said to be *well-posed* if it satisfies all of the above three properties. (Clearly, one can qualify well-posedness as being local or global in time, or being subject to some regularity condition, etc.)

For linear equations these questions are fairly simple to answer, but they become more subtle for non-linear equations. In the following we shall focus on the *nonlinear Schrödinger equation (NLS)*, a prime example for a *dispersive equation*, i.e., irregularities of solutions do not go away at all, but instead they propagate around in space. In particular, different frequency irregularities move in different directions or at different speeds. As such, solutions do not get smoother as time goes by, but they do tend to spread out and decay.

For this discussion, we shall just focus on variants of the Cauchy problem for the free linear Schrödinger equation

$$\begin{aligned} i\partial_t u - \Delta u &= 0 \\ u(0) &= f \end{aligned}$$

in two spatial dimensions, i.e., $u(t,x)$ is a function on $\mathbb{R} \times \mathbb{R}^2$. For this equation, we have the exact solution

$$u(t,x) = \mathrm{e}^{-it\Delta} f(x) = \frac{1}{4\pi t} \int_{\mathbb{R}^2} \mathrm{e}^{-i|x-y|^2/(4t)} f(y)\, dy$$

which is valid for all $t \neq 0$ (and pointwise for $f \in L^1 \cap L^2$). For fixed $f \in L^1$, the solution decays in time, namely

$$\|\mathrm{e}^{-it\Delta} f\|_\infty \leq \frac{1}{4\pi t} \|f\|_1\,.$$

On the other hand,

$$\|\mathrm{e}^{-it\Delta} f\|_2 = \|f\|_2$$

by Plancherel. (In fact, all Sobolev norms are conserved in time.) So, even though the solution decays pointwise, the L^2-norm is not altered, i.e., nothing gets "annihilated" (or created). This is reflecting the dispersive rather than dissipative (meaning that singularities attenuate and disappear as time goes by) nature of the equation. Very poetically speaking, it is the Δ term in the Schrödinger equation that causes the dispersion; without this term, the equation becomes $\partial_t u = 0$ which obviously has no dispersion.

Now let us perturb the free Schrödinger equation. Popular examples of perturbations include

- restricting the equation on a manifold on $\mathbb{R}^d$,
- adding an obstacle (and providing some sort of boundary condition),
- adding a potential,
- coupling it with another equation, or
- adding a non-linearity.

Let us consider the last option and restrict ourselves to the so-called *meson equation* or *cubic non-linear Schrödinger equation* in $d = 2$ dimensions, i.e.,

$$\begin{aligned} i\partial_t u - \Delta u &= \lambda |u|^2 u \\ u(0) &= f \end{aligned} \tag{13.2}$$

where $\lambda \in \mathbb{C}$ is a constant. One could of course consider other non-linearities as well but the cubic non-linearity is just L^2-*critical*, i.e., if $\|u\|_2$ is kept constant (as it physically is), it is not possible to scale λ away (which is possible for other powers of the non-linearity).

To get some idea of what this equation is doing, let us pretend that the dispersive term, i.e., Δu, was not present. Then (13.2) can be integrated and the solution reads

$$u(t) = (|u(0)|^{-2} - 2at)^{-\frac{a+ib}{2a}} \frac{u(0)}{|u(0)|}$$

where $-i\lambda = a + ib$. Obviously, if $a = \mathrm{Im}(\lambda) > 0$, the equation will blow up in finite time, namely at $t = (2a|u(0)|^2)^{-1}$. It is basically the non-linearity which causes a positive feedback loop and leads to the rapid increase of the solution.

However, we expect that the dispersive Δu term tries to stop this blow-up from happening by spreading the singularities of u around as soon as they get too large. Of course, for this to happen, the solution at $t = 0$ or the coupling constant λ of the non-linearity must not be too big. In fact, we have the following

Theorem 13.3. *Suppose $\|f\|_2 = 1$. Then, if λ is sufficiently small, there* ***there exists a global solution*** *u to (13.2) such that $\|u\|_2 \lesssim 1$ for all t. Furthermore, the* ***space-time estimate*** *$\|u\|_{L^4_{x,t}} \lesssim 1$ holds. This solution is* ***unique*** *subject to the above condition, and the solution* ***depends continuously*** *in the norms just mentioned on the initial data f in L^2. Finally, we have* ***scattering*** *in the following sense. There exists some initial data f_+ such that*

$$\|u(t) - \mathrm{e}^{-it\Delta} f_+\|_2 \to 0 \quad \text{as } t \to \infty\,.$$

In the PDE jargon, we just claimed that the meson equation is globally well-posed with scattering in L^2 for small λ. In Subsection 13.2.4 we shall discuss the case of large coupling constants λ.

Note that this theorem does not care about the *sign* of λ. (Intuitively, $\lambda > 0$ should act as an "attractor".) The theorem says that the decay inherent in the Schrödinger equation has tamed the effect of the non-linearity. In fact, as time goes to infinity, the non-linearity becomes increasingly irrelevant.

The techniques used to prove this result are by no means restricted to this one particular equation; they can be extended to all kinds of non-linear dispersive equations which are in some sense a small perturbation of a well-understood linear equation. Unfortunately, we still do not really understand how to push the well-posedness theory *beyond* perturbation theory into equations that are far more non-linear than (13.2).

The treatment of these equations is connected, spiritually at least, with restriction theory. An informal link is as follows. Suppose that u is a global solution to the free Schrödinger equation

$$i\partial_t u - \Delta u = 0\,.$$

Assuming that u has a space-time Fourier transform, we get (formally at least),

$$-2\pi\tau\hat{u}(\tau,\xi) + 4\pi\xi^2\hat{u}(\tau,\xi) = 0\,,$$

where

$$\hat{u}(\tau,\xi) = \int\int \mathrm{e}^{-2\pi i(t\tau+\xi\cdot x)}u(t,x)\,dt\,dx\,.$$

If $\hat{u}(\tau,\xi) \neq 0$, this implies $\tau = 2\pi|\xi|^2$, i.e., $\hat{u}$ is supported on the paraboloid

$$S = \{(\tau,\xi) : \tau = 2\pi|\xi|^2\}\,.$$

Thus, we may write

$$\hat{u} = g\ d\sigma$$

for some g, where $d\sigma$ is some surface measure on S. It turns out that the best choice of $d\sigma$ is the spatial Lebesgue measure $d\xi$, or more precisely the pullback of this measure under the projection map $(\tau,\xi) \mapsto \xi$.

If we require that the initial data of u is in L^2, it turns out to imply a L^2 estimate on g by Plancherel's theorem. In other words, we have a representation of u as $u = \widehat{gd\sigma}$ where we have L^2 control on g.

We would like to say that u decays at infinity, so that the non-linear effects will also die away. It turns out that the right estimate to use is

$$\|\widehat{gd\sigma}\|_{L^4_{x,t}} \lesssim \|g\|_2\,.$$

(In d spatial dimensions, this is $\|\widehat{gd\sigma}\|_{L^{2(d+2)/d}_{x,t}} \lesssim \|g\|_2$.) If we take the adjoint of this, this becomes

$$\|\hat{f}\|_{L^2(S,d\sigma)} \lesssim \|f\|_{4/3}$$

which is just the Tomas–Stein restriction estimate $R_S(4/3 \to 2)$ in $d = 3$ dimensions.

In what follows we will however not invoke the Tomas–Stein estimate since u is supposed to solve the *non-linear*, rather than the free Schrödinger equation. However, the Tomas–Stein philosophy, particularly squaring an estimate and interpolating between an $L^1 \to L^\infty$ and an $L^2 \to L^2$ estimate, will be very present.

13.2.2. *Proof of well-posedness in $d = 2$.* Let us start with the proof of the main theorem. We want to solve the equation

$$iu_t - \Delta u = \lambda|u|^2u$$

with initial data $u(0) = f$. Without loss of generality, let $\|f\|_2 = 1$. In a first step, we shall rewrite this equation as an integral formulation via Duhamel's principle, namely

$$u(t) = \mathrm{e}^{-it\Delta}f + i\lambda\int_0^t \mathrm{e}^{-i(t-s)\Delta}(|u|^2u)(s)\,ds\,. \tag{13.3}$$

(One should think of the first term as the influence of the initial data, whereas the second term corresponds to the cumulative influence of the forcing term $|u|^2u$.) Although this equation is equivalent to the differential form, it is much easier to handle when it comes to proving existence and uniqueness.

To find a solution, we shall use an iterative method. We first approximate u by the linear solution

$$u_0(t) = \mathrm{e}^{-it\Delta} f$$

and then make the better approximation

$$u_1(t) = \mathrm{e}^{-it\Delta} f + i\lambda \int_0^t \mathrm{e}^{-i(t-s)\Delta}(|u_0|^2 u_0)\, ds$$

and so forth, by defining $u_{k+1} = Nu_k$ where

$$(N(u))(t) = \mathrm{e}^{-it\Delta} f + i\lambda \int_0^t \mathrm{e}^{-i(t-s)\Delta}(|u|^2 u)\, ds\,.$$

We hope that this sequence of approximations converges to a limit as $k \to \infty$ so that $N(u) = u$. Put differently, our goal is to show that the operator N has a fixed point, that this point is unique, and that it depends continuously on the data. This would be an immediate consequence of the contraction mapping theorem, provided we know that N is a contraction on some metric space X which contains u_0. This sounds easy enough – and a very large number of existence results in PDE are ultimately derived from this very simple idea. The catch is that we have to pick the right metric space to get the contraction working.

After a lot of experimentation and looking at the behavior of the first few iterates u_0, u_1, etc., we ultimately decide that the correct space to use is

$$X = \{u : \|u\|_{L^4_{x,t}} \le C\}\,,$$

where C is some universal constant and the metric is induced by the L^4 norm. Thus, we would like to show

$$\|u_0\|_{L^4_{x,t}} \lesssim 1 \tag{13.4}$$

and

$$\|N(u) - N(v)\|_{L^4_{x,t}} \le \frac{1}{2}\|u - v\|_{L^4_{x,t}} \quad \text{for all } u, v \in L^4_{x,t}\,. \tag{13.5}$$

This will be accomplished by the following three estimates which go under the name Strichartz estimates and were already discussed in Subsection 13.1. We shall use the homogeneous Strichartz estimate (yields estimate (13.4) on $u_0 = \mathrm{e}^{-it\Delta} f$)

$$\|\mathrm{e}^{-it\Delta} f\|_{L^4_{x,t}} \lesssim \|f\|_2 \tag{13.6}$$

(compare this to Theorem 13.1), the dual homogeneous Strichartz estimate (yields scattering and that $u(t)$ still belongs to L^2)

$$\|\int_0^\infty \mathrm{e}^{it\Delta} F(t)\, dt\|_2 \lesssim \|F\|_{L^{4/3}_{x,t}}\,, \tag{13.7}$$

and the retarded Strichartz estimate (yields the contraction property)

$$\|\int_0^t \mathrm{e}^{-i(t-s)\Delta} F(s)\, ds\|_{L^4_{x,t}} \lesssim \|F\|_{L^{4/3}_{x,t}}\,. \tag{13.8}$$

We shall prove these estimates in the next subsubsection. For now, let us see how these estimates give what we want.

First, the estimate on the zeroth iteration $u_0 = \mathrm{e}^{-it\Delta} f$, i.e., (13.4), follows trivially from (13.6). Next, we shall prove the contraction property (13.5). We first note that one can simplify $N(u) - N(v)$ as

$$N(u) - N(v) = i\lambda \int_0^t \mathrm{e}^{-i(t-s)\Delta}(|u|^2 u - |v|^2 v)\, ds\,.$$

Thus, by (13.8), we have

$$\|N(u)-N(v)\|_{L^4_{x,t}} \lesssim |\lambda| \||u|^2u-|v|^2v\|_{L^{4/3}_{x,t}}\,.$$

Now, we use the pointwise estimate

$$\begin{aligned}||u|^2u-|v|^2v| &= ||u|^2(u-v)-|v|^2(v-u)+|u|^2v-|v|^2u| \\ &\le ||u|^2(u-v)-|v|^2(v-u)|+|u\overline{u}v-v\overline{v}u| \\ &= ||u|^2(u-v)-|v|^2(v-u)|+|uv(\overline{u}-\overline{v})| \\ &\le |u|^2|u-v|+|v|^2|v-u|+|u||v||u-v| \\ &\le \frac{3}{2}\left[|u|^2|u-v|+|v|^2|u-v|\right] = \mathcal{O}(|u|^2|u-v|)+\mathcal{O}(|v|^2|u-v|)\end{aligned}$$

and Hölder's inequality to obtain

$$\|N(u)-N(v)\|_{L^4_{x,t}} \lesssim |\lambda| \left(\|u\|_4^2\|u-v\|_4+\|v\|_4^2\|u-v\|_4\right)\,.$$

But since $u,v\in L^4_{x,t}$, (13.5) clearly holds, if λ is chosen sufficiently small.

Thus, we have proven existence, uniqueness, and continuous dependence of u on the initial data. As a bonus, we get that the limit $u\in L^4_{x,t}$. However, we are not done yet; we still need to show that $u(t)$ still belongs to L^2 and that scattering occurs. Let us investigate the square-integrability. From Duhamel's version of the NLS, we obtain

$$\|u(t)\|_2 \lesssim \|\mathrm{e}^{-it\Delta}f\|_2+\|\mathrm{e}^{-it\Delta}\int_0^t \mathrm{e}^{is\Delta}(|u|^2u)\,ds\|_2\,.$$

Clearly, the first term is bounded since $f\in L^2$. To estimate the second one, we use (13.7) and obtain

$$\|\int_0^t \mathrm{e}^{is\Delta}(|u|^2u)\,ds\|_2 \lesssim \||u|^2u\|_{4/3} = \|u\|_4^3 \lesssim 1$$

as desired. Finally, we show scattering. Define f_+ by

$$f_+ = f+i\lambda\int_0^\infty \mathrm{e}^{is\Delta}(|u|^2u)(s)\,ds\,,$$

i.e., f_+ is equal f modified by the backdated effect of the non-linearity. From (13.7) and the argument just given, we see that $f_+\in L^2$. We wish to show

$$\|u(t)-\mathrm{e}^{-it\Delta}f_+\|_2\to 0 \quad \text{as } t\to\infty\,.$$

From Duhamel's version of the NLS, we have

$$\begin{aligned}u(t)-\mathrm{e}^{-it\Delta}f_+ &= \left(\mathrm{e}^{-it\Delta}f+i\lambda\int_0^t \mathrm{e}^{-i(t-s)\Delta}(|u|^2u)(s)\,ds\right) \\ &\quad -\left(\mathrm{e}^{-it\Delta}f+i\lambda\int_0^\infty \mathrm{e}^{-i(t-s)\Delta}(|u|^2u)(s)\,ds\right) \\ &= -i\lambda\mathrm{e}^{-it\Delta}\int_t^\infty \mathrm{e}^{is\Delta}(|u|^2u)(s)\,ds\,.\end{aligned}$$

Using (13.7), we obtain

$$\|u(t)-\mathrm{e}^{-it\Delta}f_+\|_2 \lesssim |\lambda|\|\mathbf{1}_{[t,\infty)}|u|^2u\|_{4/3}$$

which yields the claim by monotone convergence.

13.2.3. *Proof of the Strichartz estimates.* `[Check whether the following arguments were generalized by Keel--Tao [119, Theorem 1.2] to obtain sharp Strichartz estimates from` $L^1 \to L^\infty$ `bounds on` e^{itH}`.]`

Let us first see the implications (13.8) ⇒ (13.7) ⇒ (13.6) and finally prove (13.8). First, the homogeneous Strichartz estimate follows from the dual homogeneous estimate by Cauchy–Schwarz, namely

$$\left|\int \overline{\left(\int_0^\infty \mathrm{e}^{it\Delta} F(t,x)\,dt\right)} f(x)\,dx\right| \lesssim \|F\|_{L^{4/3}_{x,t}} \|f\|_2 \,.$$

Rearranging the left side, this becomes

$$\left|\int\int \overline{F(t,x)}\mathrm{e}^{-it\Delta} f(x)\,dt\,dx\right| \lesssim \|F\|_{L^{4/3}_{x,t}} \|f\|_2 \,.$$

Taking $\sup_{F\in L^{4/3}_{x,t}}$, we obtain $\|\mathrm{e}^{-it\Delta} f\|_{L^4_{x,t}} \lesssim \|f\|_2$, i.e., (13.6) by duality of the L^p spaces.

Next, let us see how the dual homogeneous estimate follows from the retarded estimate. First, we square the dual homogeneous estimate as

$$\langle \int_0^\infty \mathrm{e}^{it\Delta} F(t)\,dt, \int_0^\infty \mathrm{e}^{is\Delta} F(s)\,ds\rangle \lesssim \|F\|^2_{L^{4/3}_{x,t}}$$

and rewrite it as

$$\int_0^\infty \int_0^\infty \langle \mathrm{e}^{it\Delta} F(t), \mathrm{e}^{is\Delta} F(s)\rangle\,ds\,dt \lesssim \|F\|^2_{L^{4/3}_{x,t}} \,.$$

By symmetry, it suffices to consider the portion of the double integral where $s \le t$, i.e.,

$$\int_0^\infty dt \int_0^t ds\ \langle \mathrm{e}^{it\Delta} F(t), \mathrm{e}^{is\Delta} F(s)\rangle \lesssim \|F\|^2_{L^{4/3}_{x,t}}$$

and rewrite this once more as

$$\int_{\mathbb{R}^2} \int_0^\infty F(t,x) \left(\int_0^t \mathrm{e}^{-i(t-s)\Delta} F(s,x)\,ds\right)\,dt\,dx \,.$$

Now, by Hölder's inequality, the left side is bounded by

$$\|F\|_{L^{4/3}_{x,t}} \left\|\int_0^t \mathrm{e}^{-i(t-s)\Delta} F(s,x)\,ds\right\|_{L^4_{x,t}} .$$

Now, we may apply the retarded estimate (13.8) which yields the claimed inequality.

Finally, let us prove the retarded estimate. We right out the $L^4_{x,t}$ norm of the left side of (13.8) as

$$\left(\int_0^\infty \mathrm{d}t \left\|\int_0^t \mathrm{e}^{-i(t-s)\Delta} F(s)\,ds\right\|^4_{L^4_x}\right)^{1/4}$$

and apply Minkowski's inequality to obtain

$$\| \int_0^t \mathrm{e}^{-i(t-s)\Delta} F\,ds\|_{L^4_{x,t}} \le \left(\int_0^\infty dt \left(\int_0^t ds\ \|\mathrm{e}^{-i(t-s)\Delta} F\|_{L^4_x}\right)^4\right)^{1/4}$$

By interpolating between $\|e^{-it\Delta}f\|_2 \leq \|f\|_2$ and $\|e^{-it\Delta}f\|_\infty \lesssim |t|^{-1}\|f\|_1$, we can estimate the L_x^4 norm appearing in the integrand and thus obtain

$$\begin{aligned}\|\int_0^t e^{-i(t-s)\Delta}F\,ds\|_{L^4_{x,t}} &\lesssim \left(\int_0^\infty dt\left(\int_0^t ds\ |t-s|^{-1/2}\|F(s)\|_{L_x^{4/3}}\right)^4\right)^{1/4}\\ &\leq \||\cdot|^{-1/2} * \|F(\cdot)\|_{L_x^{4/3}}\|_{L_t^4} \lesssim \|F\|_{L^{4/3}_{x,t}}\end{aligned}$$

where we used the Hardy–Littlewood–Sobolev inequality (in the time-dimension!) in the final inequality. This concludes the proof of (13.8).

13.2.4. *Large values of λ.* The proof of the contraction property (13.5) crucially relied on the smallness of λ. Therefore, it is not expected to obtain global existence for large λ since the non-linear term can make the the wave function extremely large for certain frequencies before dispersive effect of Δ can repair the damage. However, one can at least get a local solution.

Theorem 13.4. *Suppose $\|f\|_2 = 1$ and λ is arbitrary. Then there exits a time $T_0 > 0$ and a local solution u to* (13.2) *such that $\|u(t)\|_2 \lesssim 1$ for all $0 \leq t \leq T_0$. Moreover, u satisfies $\|u\|_{L^4_{x,t}(\mathbb{R}^2\times[0,T_0])} \lesssim 1$. This solution u is unique subject to the above conditions and the solution depends continuously in the norms just mentioned on the initial data $f \in L^2$.*

The proof of this theorem is virtually identical to that of the main theorem. There are, however, two main differences.

(1) All our norms are restricted to the space-time interval $\mathbb{R}^3 \times [0, T_0]$.
(2) We iterate on a much smaller ball, namely
$$X = \{u : \|u\|_{L^4_{x,t}(\mathbb{R}^2\times[0,T_0])} \leq \varepsilon\}$$
where $\varepsilon(\lambda)$ is a tiny number. One can check that the Duhamel map N is still a contraction if this number is small enough.
(3) We must guarantee that the zeroth iterate $u_0 = e^{-it\Delta}f$ is in X. But this follows from the homogeneous Strichartz estimate $\|e^{-it\Delta}f\|_{L^4_{x,t}} \lesssim \|f\|_2$, i.e., $\|u_0\|_{L^4_{x,t}} \lesssim 1$ globally. Thus, if T_0 is chosen small enough, monotone convergence shows $\|u_0\|_{L^4_{x,t}(\mathbb{R}^2\times[0,T_0])} \leq \varepsilon$.

13.3. **Strichartz estimates for the Schrödinger equation on the torus via decoupling inequalities.** See Subsubsection 21.1.2 and the notes of Hickman–Vitturi [107, p. 22, Lecture 2, Section 2.2].

14. Pointwise convergence of the Schrödinger evolution

We consider the nonlinear Schrödinger equation (NLS)

$$\begin{cases} i\partial_t u(x,t) = -\Delta u(x,t) + \mathcal{N}(u)\,, \\ u(x,0) = f(x)\,, \end{cases} \qquad x \in \mathbb{R}^d \text{ or } \mathbb{T}^d$$

where $\mathbb{T}^d = \mathbb{R}/2\pi\mathbb{Z}$ and $\mathcal{N}$ is a power-type nonlinearity. The basic question is the following: Let $s > 0$ and $f \in H^s(\mathbb{R}^d)$. For which $s > 0$ does the solution $u(x,t)$ converge pointwise (Lebesgue) almost everywhere to $f(x)$ as $t \to 0$? For $\mathcal{N} = 0$ in $\mathbb{R}^1$ this question was first posed by Carleson [43, p. 24] who showed that almost everywhere convergence holds, whenever $f \in H^{1/4}(\mathbb{R})$. Dahlberg–Kenig [60] showed that this one-dimensional result is sharp; in fact, they proved that $s \geq 1/4$ is a necessary condition for a.e. convergence in $\mathbb{R}^d$ for all $d \geq 1$. Recently, Bourgain [23] showed that $s \geq d/(2(d+1))$ is a necessary condition for a.e. convergence to the initial data. This has been proved to be sharp, up to the endpoint, by Du–Guth–Li [65] in $d = 2$ and Du–Zhang [67] in higher dimensions.

See works by Kenig–Ruiz [121], Sjölin [157], Vega [192], Bourgain [17, 23], Du–Guth–Li [65], Du–Guth–Li–Zhang [66], Du–Zhang [67], and the references therein.

For $\mathcal{N}(z) = |z|^{p-1}z$, see Compaan–Lucá–Staffilani [48] who proved pointwise a.e. convergence in $\Omega^d \in \{\mathbb{R}^d, \mathbb{T}^d\}$ for $p \geq 3$ and

$$s > \max\left\{s^*, \frac{d}{2} - \frac{2}{p-1}\right\}$$

and $s \geq 1/4$ for $d = 1$ and $p < 9$. Here,

$$s^* := \inf\left\{s : \lim_{t\to 0} \mathrm{e}^{it\Delta} f(x) = f(x) \text{ for a.e. } x \in \Omega^d\,,\ f \in H^s(\Omega^d)\right\}$$

is the exponent for which pointwise a.e. convergence in the linear setting, i.e., $\mathcal{N} = 0$ holds, i.e., $s^* = d/(2(d+1))$ on $\mathbb{R}^d$ and $s^* = d/(d+2)$ on $\mathbb{T}^d$ (see Moyua–Vega [140] in $d = 1$ and Wang–Zhang [195] for higher dimension.).

See Dimou–Seeger [64] for convergence of evolution generated by fractional Laplace in one dimension. Bounds on eigenfunctions of $-\Delta$, see, e.g., Sogge [160, 159]. See also Stovall's review [171] for more references.

15. Connection to the Kakeya conjecture

In this section, we first show that the so-called Kakeya maximal conjecture is a consequence of the restriction conjecture. Afterwards, we discuss the connection between the so-called Kakeya set conjecture and the Kakeya maximal conjecture. In particular, we review the proof of the two-dimensional Kakeya maximal conjecture. Finally, we discuss how Kakeya can be used to study the restriction conjecture without the help of the square function conjecture (see Section 8).

15.1. **Restriction conjecture ⇒ Kakeya maximal conjecture.** Here, we follow Lecture 1 in the notes of Hickman and Vitturi [107] and Wolff [200, Proposition 10.5].

The Knapp example (Subsection 3.2) in the introduction is central to the following discussion. Recall that the restriction estimate $R^*_{\mathbb{P}^{d-1}}(q' \to p')$ "just barely fails" for $p' = q' = 2d/(d-1)$. By that we mean that for all $\varepsilon > 0$ the estimate

$$\|(Fd\sigma)^\vee\|_{L^{2d/(d-1)}(B(0,R))} \lesssim R^\varepsilon \|F\|_{L^{2d/(d-1)}(\mathbb{P}^{d-1},d\sigma)}$$

holds for all $R \gg 1$.

We are now going to consider the case where F is the superposition of many disjoint Knapp examples, i.e.,

$$F = \sum_\kappa \mathbf{1}_\kappa$$

where κ is a $R^{-1/2} \times \cdots \times R^{-1/2}$ cap on $\mathbb{P}^{d-1}$. If $\Omega \subseteq \mathbb{S}^{d-1}$ denotes the set of normal directions to these caps, we have

$$\|F\|_{L^{2d/(d-1)}(\mathbb{P}^{d-1})} \lesssim \left(R^{-(d-1)/2} \times |\Omega|\right)^{(d-1)/(2d)} . \tag{15.1}$$

On the other hand, the uncertainty principle tells us that $(\mathbf{1}_\kappa d\sigma)^\vee$ is essentially constant on a tube dual to κ (with unit normal ω), i.e., on a tube T_ω with dimensions $R^{1/2} \times \cdots \times R^{1/2} \times R$ which is oriented in the direction ω. Away from T_ω, the function $(\mathbf{1}_\kappa d\sigma)^\vee$ decays rapidly. Thus, heuristically,

$$(Fd\sigma)^\vee(x) \sim R^{-(d-1)/2} \sum_{\omega\in\Omega} \mathrm{e}^{2\pi i x\cdot\xi_\omega} \mathbf{1}_{T_\omega}(x) \tag{15.2}$$

where ξ_ω denotes the center of the cap oriented in direction $\omega \in \Omega$. By modulating the summands of F, one may replace each T_ω in (15.2) with any translate of itself while maintaining (15.1). Here, we will however agree that the tubes are contained in a ball $B(0, AR)$ for some $A > 1$ but otherwise arrange them in an arbitrary fashion. Our goal now is to show that these tubes are "essentially disjoint" even if their overlap is "maximal" (which it is if we translate them in the above fashion).

Due to the summation over exponentials in (15.2), we expect considerable cancellations. If no cancellation was present, $|(Fd\sigma)^\vee|$ would roughly equal $R^{-(d-1)/2}$ times the ℓ^1 sum of the $\mathbf{1}_{T_\omega}$. Because of the cancellations, we expect that the ℓ^1 sum should be replaced by a smaller ℓ^2 sum. In fact, randomizing this sequence lets us exploit these cancellations effectively via Khintchine's inequality (see, e.g., Stein [165, Chapter IV, §5, Equation (44) and Appendix D]), i.e.,

$$\|(\sum_k |g_k|^2)^{1/2}\|_p^p \sim \int_{\mathbb{R}^d} \mathbb{E}\{|\sum_k \varepsilon_k g_k(x)|^p\},$$

where $(\varepsilon_k)_k$ is a Rademacher distributed sequence, i.e., a sequence of statistically independent and identically distributed random variables with $P(\varepsilon_k = \pm 1) = 1/2$ for all k.

Thus, instead of considering a mere sum of Knapp examples, we define the modulated and randomized sum

$$F(\xi) = \sum_\kappa \varepsilon_\kappa \mathrm{e}^{2\pi i x_\kappa \cdot \xi} \mathbf{1}_\kappa(\xi)$$

for some choice of $x_\kappa \in \mathbb{R}^d$. Note that

$$\begin{aligned} \mathbb{E}\|(Fd\sigma)^\vee\|_{L^{\frac{2d}{d-1}}(\mathbb{R}^d)}^{\frac{2d}{d-1}} &\lesssim R^\varepsilon \mathbb{E}\|F\|_{L^{\frac{2d}{d-1}}(\mathbb{P}^{d-1})}^{\frac{2d}{d-1}} \\ &= R^\varepsilon \int_{\mathbb{P}^{d-1}} \mathbb{E}\{|\sum_\kappa \varepsilon_\kappa \mathrm{e}^{2\pi i x_\kappa \cdot \xi} \mathbf{1}_\kappa(\xi)|^{\frac{2d}{d-1}}\} \sim R^{\varepsilon - (d-1)/2} \cdot |\Omega| \end{aligned} \tag{15.3}$$

by the restriction conjecture and since the value of $|F|$ is independent of the outcome of the ε_κ. Moreover,

$$(Fd\sigma)^\vee(x) = \sum_\kappa \varepsilon_\kappa (\mathbf{1}_\kappa d\sigma)^\vee(x - x_\kappa).$$

Applying Khintchine's inequality and the uncertainty principle (15.2), we obtain

$$\begin{aligned} \mathbb{E}\|(Fd\sigma)^\vee\|_{L^{\frac{2d}{d-1}}(\mathbb{R}^d)}^{\frac{2d}{d-1}} &\sim \|(\sum_\kappa |(\mathbf{1}_\kappa d\sigma)^\vee(\cdot - x_\kappa)|^2)^{1/2}\|_{\frac{2d}{d-1}}^{\frac{2d}{d-1}} \gtrsim \|R^{-(d-1)/2}(\sum_\omega \mathbf{1}_{T_\omega})^{1/2}\|_{\frac{2d}{d-1}}^{\frac{2d}{d-1}} \\ &= R^{-d}\|\sum_\omega \mathbf{1}_{T_\omega}\|_{\frac{d}{d-1}}^{\frac{d}{d-1}} \end{aligned}$$

Combining this with (15.3), we obtain (noting $|T_\omega|^{\frac{d-1}{d}} \sim R^{\frac{(d+1)(d-1)}{2d}} = R^{(d-1)-\frac{(d-1)^2}{2d}}$)

$$\|\sum_\omega \mathbf{1}_{T_\omega}\|_{\frac{d}{d-1}} \lesssim R^{\varepsilon + (d-1) - \frac{(d-1)^2}{2d}} \times |\Omega|^{\frac{d-1}{d}} = R^\varepsilon \left(|T_\omega| \cdot |\Omega|\right)^{\frac{d-1}{d}}.$$

We may thus summarize our above findings in the following conjecture which would follow from the restriction conjecture.

Conjecture 15.1 (Kakaya maximal conjecture)**.** *Let $\Omega \subseteq \mathbb{S}^{d-1}$ be a maximal set of $R^{-1/2}$-separated directions and $(T_\omega)_{\omega \in \Omega}$ a collection of $R^{1/2} \times \cdots \times R^{1/2} \times R$-rectangles where T_ω is* *oriented in the direction of* ω*. Then, for any* $\varepsilon > 0$*, the inequality*

$$\Big\|\sum_{\omega\in\Omega}\mathbf{1}_{T_\omega}\Big\|_{L^{d/(d-1)}(\mathbb{R}^d)} \lesssim_\varepsilon R^\varepsilon\left(\sum_{\omega\in\Omega}|T_\omega|\right)^{\frac{d-1}{d}} \tag{15.4}$$

holds.

The reason why the above conjecture is called a maximal conjecture is that it can be reformulated in terms of one, as we shall see in a moment, but see in particular Conjecture 15.2 and Lemma 15.6. (In short, Lemma 15.6 says that Conjecture 15.1 implies Conjecture 15.2 which is the maximal Kakeya conjecture stated in the usual form.)

Let us continue with the discussion of Conjecture 15.1. If the rectangles T_ω were mutually disjoint, the above inequality (15.4) would of course be an equality, i.e.,

$$\Big\|\sum_{\omega\in\Omega}\mathbf{1}_{T_\omega}\Big\|_{L^{d/(d-1)}(\mathbb{R}^d)} = R^\varepsilon\left(\sum_{\omega\in\Omega}|T_\omega|\right)^{\frac{d-1}{d}}.$$

This means that (15.4) can be interpreted as the statement that the rectangles pointing in different directions must have small intersection, i.e., they must be "essentially disjoint". This heuristic can in fact be made more precise. Let us define the overlap of the tubes T_ω by α, i.e.,

$$|\bigcup_\omega T_\omega| = \alpha\sum_{\omega\in\Omega}\mathbf{1}_{T_\omega}.$$

Clearly, $0 < \alpha \leq 1$. From Hölder's inequality, we have

$$\sum_{\omega\in\Omega}|T_\omega| = \|\sum_{\omega\in\Omega}\mathbf{1}_{T_\omega}\|_1 \leq \|\sum_{\omega\in\Omega}\mathbf{1}_{T_\omega}\|_{d/(d-1)}|\bigcup_{\omega\in\Omega}T_\omega|^{1/d}.$$

Combining this with (15.4), we obtain

$$R^{-\varepsilon} \leq \alpha,$$

i.e., α is essentially 1 up to extremely small powers of δ.

Let us now finally explain, why the name "maximal conjecture" is appropriate. In the following, we assume $0 < \delta \ll 1$ and f be a compactly supported function and define the Kakeya maximal function by

$$f_\delta^*(\omega) := \sup_{T_\omega}\frac{1}{|T_\omega|}\int_{T_\omega}|f| \tag{15.5}$$

where the supremum is taken over all $1\times\delta\times\cdots\times\delta$ tubes T_ω which are oriented in the direction of $\omega\in\mathbb{S}^{d-1}$. Let $K(p,\varepsilon)$ denote the estimate

$$\|f_\delta^*\|_{L^p(\mathbb{S}^{d-1})} \lesssim \delta^{-d/p+1-\varepsilon}\|f\|_{L^p(\mathbb{R}^d)}. \tag{15.6}$$

Conjecture 15.2 (Kakaya maximal conjecture – equivalent formulation)**.** *Let* $0 < \delta \ll 1$*,* f*, and* f_δ^* *be as above. Then* $K(p,\varepsilon)$ *holds for all* $1 \leq p \leq d$ *and* $\varepsilon > 0$*.*

Bourgain [15, Section 2] proved this conjecture, provided $1 \leq p \leq p(d)$ where

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

i.e., in particular $(d+1)/2 < p(d) < (d+2)/2$. Below, we shall focus on the case $p = d$ which corresponds to a "$\delta^{-\varepsilon}$-estimate" in the above conjecture. Some remarks on Conjecture 15.2 are in order.

Remarks 15.3. (1) It is clear from the definition that

$$\|f_\delta^*\|_\infty \leq \|f\|_\infty\,, \tag{15.7}$$

$$\|f_\delta^*\|_\infty \leq \delta^{-(d-1)}\|f\|_1\,. \tag{15.8}$$

(2) If $d \geq 2$ and $p < \infty$, then there can be no bounds of the form

$$\|f_\delta^*\|_q \leq C\|f\|_p\,, \tag{15.9}$$

where C is independent of δ. (The role of q is not important here.) To see this, consider a zero measure Kakeya set E [14], let E_δ be the δ-neighborhood of E, and $f := \mathbf{1}_{E_\delta}$. Then $f_\delta^*(\omega) = 1$ for all $\omega \in \mathbb{S}^{d-1}$ and hence $\|f_\delta^*\|_q \sim 1$. But on the other hand, $\lim_{\delta\to 0}\|f\|_p^p = \lim_{\delta\to 0}|E_\delta| = 0$ for any $p < \infty$.

(3) Let $f = \mathbf{1}_{B_0(\delta)}$. Then for all $\omega \in \mathbb{S}^{d-1}$ the tube $T_\omega^\delta(0)$ contains $B_0(\delta)$ so that

$$f_\delta^*(\omega) = \frac{|B_0(\delta)|}{|T_\omega^\delta(0)|} \gtrsim \delta\,.$$

Hence, $\|f_\delta^*\|_p \sim \delta$. But on the other hand, we have $\|f\|_p \sim \delta^{d/p}$ which ultimately shows that a "$\delta^{-\varepsilon}$-estimate" of the form

$$\forall\varepsilon > 0\ \exists C_\varepsilon > 0:\ \|f_\delta^*\|_{L^p(\mathbb{S}^{d-1})} \leq C_\varepsilon\delta^{-\varepsilon}\|f\|_{L^p(\mathbb{R}^d)}$$

can never hold for any $p < d$. Thus, by interpolation with (15.7), the Kakeya problem therefore consists in establishing

$$\forall\varepsilon\ \exists C_\varepsilon:\ \|f_\delta^*\|_{L^d(\mathbb{S}^{d-1})} \leq C_\varepsilon\delta^{-\varepsilon}\|f\|_{L^d(\mathbb{R}^d)}\,. \tag{15.10}$$

In fact, this was proved for $d = 2$ by Córdoba [51] in a somewhat different formulation and by Bourgain [15] as we stated it here. These results are somewhat easy in $d = 2$ since the L^2-formalism (with all its measures of orthogonality and oscillations through Plancherel) can be exploited heavily.

(4) Interpolating (15.10) with (15.8) on the other hand gives a family of conjectured inequalities

$$\|f_\delta^*\|_{L^q(\mathbb{S}^{d-1})} \lesssim_\varepsilon \delta^{-d/p+1-\varepsilon}\|f\|_{L^p(\mathbb{R}^d)}\,, \quad 1 \leq p \leq d \quad \text{and} \quad q = (d-1)p'\,. \tag{15.11}$$

Note that if (15.11) holds for some $p_0 > 1$, then it also holds for all $1 \leq p \leq p_0$ (by interpolation with (15.8)). The current best results in this direction are that (15.11) holds with $p = \min\{(d+2)/2, (4d+3)/7\}$ and a suitable q, see Wolff [197] and Katz–Tao [118]. In fact, Wolff established the $p(d) = (d+2)/2$ endpoint in Bourgain's result with $q = (d-1)p'$. As we shall see soon (Proposition 15.13), this implies that Kakeya sets have Hausdorff dimension $\geq (d+2)/2$ (see also Remark 15.23). In Theorem 15.19 we shall prove Bourgain's result. $\square$

Let us now give another proof of the fact that the restriction conjecture implies the Kakeya maximal conjecture.

Proposition 15.4 (Fefferman, Bourgain)**.** *Assume that the restriction estimate*

$$\|\widehat{fd\sigma}\|_p \lesssim_p \|f\|_{L^p(\mathbb{S}^{d-1})}\,, \qquad p > \frac{2d}{d-1} \tag{15.12}$$

holds. Then, the Kakeya maximal estimate (15.13)*, i.e.,*

$$\|f_\delta^*\|_{L^d(\mathbb{S}^{d-1})} \lesssim_\varepsilon \delta^{-\varepsilon}\|f\|_{L^d(\mathbb{R}^d)}$$

holds true.

[14] Such sets can be constructed explicitly, see, e.g., Besicovitch [9], Perron [145], Kahane [115].

Remark 15.5. Note that (15.12) is only ostensibly stronger than (15.33) below which states $\|\widehat{fd\sigma}\|_p \lesssim_p \|f\|_{L^\infty(\mathbb{S}^{d-1})}$ for $p > 2d/(d-1)$. In fact, these estimates are (formally at least) equivalent, see Bourgain [15].

The proof of Proposition 15.4 relies on the following Lemma, which will come in handy later in the proof of the two-dimensional Kakeya conjecture (by Córdoba [51]). In what follows, we denote by

$$T_e^\delta(a) = \{x \in \mathbb{R}^d : \ |(x-a)\cdot e| \le \frac{1}{2}, \ |(x-a)^\perp| \le \delta\}, \quad x^\perp = x - (x\cdot e)e$$

the δ-neighborhood of a unit line segment in the e direction, centered at a.

Lemma 15.6. *Let $0 < \delta \ll 1$ and $1 < p < \infty$ and suppose p has the following property: if $\{e_k\} \subseteq \mathbb{S}^{d-1}$ is a maximal δ-separated set, and if $\delta^{d-1}\sum_k y_k^{p'} \le 1$, then for any choice of points $a_k \in \mathbb{R}^d$, we have*

$$\|\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\|_{p'} \le A\,.$$

Then, there is a bound

$$\|f_\delta^*\|_{L^p(\mathbb{S}^{d-1})} \lesssim A\|f\|_{L^p(\mathbb{R}^d)}\,.$$

Remark 15.7. Observe that the maximal δ-separated subset $\{e_k\}$ of $\mathbb{S}^{d-1}$ has cardinality $\approx \delta^{-(d-1)}$.

Proof. Let $\{e_k\}_k$ be a maximal δ-separated subset of $\mathbb{S}^{d-1}$. Observe that if $|\omega-\omega'| < \delta$, then $f_\delta^*(\omega) \le C f_\delta^*(\omega')$ since any $T_\omega^\delta(a)$ can be covered by a bounded number of tubes $T^\delta_{\omega'}(a')$. Therefore,

$$\|f_\delta^*\|_p \le \left(\sum_k \int_{B_{e_k}(\delta)} |f_\delta^*(\omega)|^p\, d\omega\right)^{1/p} \le C\left(\delta^{d-1}\sum_k |f_\delta^*(e_k)|^p\right)^{1/p} = C\delta^{d-1}\sum_k y_k |f_\delta^*(e_k)|$$

for some sequence y_k with $\sum_k y_k^{p'}\delta^{d-1} = 1$, where we used the duality between ℓ^p and $\ell^{p'}$ (i.e., $\|f\|_{\ell^p} = \langle g, f\rangle$ for some $(g_k)_{k\in\mathbb{N}} \in \ell^{p'}$ with $\|g\|_{\ell^{p'}} = 1$; here, $f_k = \delta^{(d-1)/p}|f_\delta^*(e_k)|$ and $g_k = y_k\delta^{(d-1)/p')}$) in the last line. Therefore, by the definition of the maximal function,

$$\|f_\delta^*\|_p \lesssim \delta^{d-1}\sum_k y_k \frac{1}{|T^\delta_{e_k}(a_k)|}\int_{T^\delta_{e_k}(a_k)} |f|$$

for a certain choice of $\{a_k\}$. But since $|T^\delta_{e_k}(a_k)| \sim \delta^{d-1}$, we obtain (using Hölder and the hypothesis),

$$\|f_\delta^*\|_p \lesssim \int \left(\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\right)|f| \le \|\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\|_{p'}\|f\|_p \le A\|f\|_p\,,$$

thereby proving the lemma. □

Proof of Proposition 15.4. In view of the above lemma, we chose a maximal δ-separated subset $\{e_k\}$ of $\mathbb{S}^{d-1}$ whose cardinality is roughly $\delta^{-(d-1)}$, as observed before. Now, for each j, pick a tube $T^\delta_{e_j}(a_j) \equiv T_j$ and denote by τ_j the δ^{-2} rescaled version of T_j, i.e., the tube of length δ^{-2} and thickness δ^{-1}, oriented along the e_j direction. Furthermore, let

$$S_j = \{\omega \in \mathbb{S}^{d-1} : \ |1 - \omega\cdot e_j| \le C^{-1}\delta^2\}$$

be a spherical cap of radius $\sim C^{-1}\delta$, centered at e_j. Here, C is chosen so large that the S_j are pairwise disjoint. (Note that the S_j are just dual to the τ_j.) Now, let f_j be the associated Knapp examples, i.e., f_j are supported on S_j and satisfy

$$\|f_j\|_{L^\infty(\mathbb{S}^{d-1})} = 1$$
$$|\widehat{f_j d\sigma}| \gtrsim \delta^{d-1} \quad \text{on } \tau_j\,.$$

Now, let

$$f_\varepsilon := \sum_j \varepsilon_j y_j f_j\,,$$

where the y_j are non-negative weights and the sequence $\{\varepsilon_j\}_j$ is a Rademacher sequence. Since the f_j have disjoint supports, we have on the one hand

$$\|f_\varepsilon\|_{L^q(\mathbb{S}^{d-1})}^q = \sum_j y_j^q \|f_j\|_{L^q(\mathbb{S}^{d-1})}^q \sim \sum_j y_j^q \delta^{d-1}$$

since $|\mathbb{S}^{d-1} \cap S_j| \sim \delta^{d-1}$ for all j. On the other hand, we have by Khintchine's inequality

$$\|(\sum_k |g_k|^2)^{1/2}\|_p^p \sim \int_{\mathbb{R}^d} \mathbb{E}\{|\sum_k \varepsilon_k g_k(x)|^p\}\,,$$

that

$$\begin{aligned}\mathbb{E}(\|\widehat{f_\varepsilon d\sigma}\|_{L^q(\mathbb{R}^d)}^q) = \int_{\mathbb{R}^d} \mathbb{E}(|\widehat{f_\varepsilon d\sigma}(x)|^q)\,dx \sim \int_{\mathbb{R}^d} \left(\sum_j y_j^2 |\widehat{f_j d\sigma}(x)|^2\right)^{q/2} dx \\ \gtrsim \delta^{q(d-1)} \int_{\mathbb{R}^d} |\sum_j y_j^2 \mathbf{1}_{\tau_j}(x)|^{q/2}\,dx.\end{aligned}$$

Now, assuming that the restriction estimate (15.12) holds true, we can combine the last two inequalities and obtain for any $q > 2d/(d-1)$,

$$\delta^{q(d-1)} \int_{\mathbb{R}^d} |\sum_j y_j^2 \mathbf{1}_{\tau_j}|^{q/2}\,dx \lesssim \sum_j y_j^q \delta^{d-1}\,.$$

This is almost the estimate that we need to apply Lemma 15.6. Introducing $z_j = y_j^2$ and $p' = q/2$, the above inequality is equivalent to the statement

$$\text{if } \delta^{d-1} \sum_j z_j^{p'} \leq 1\,, \text{then } \|\sum_j z_j \mathbf{1}_{\tau_j}\|_{p'} \lesssim \delta^{-2(d-1)}$$

for any $p' \geq d/(d-1)$. Now, rescaling by δ^2, the above is equivalent to

$$\text{if } \delta^{d-1} \sum_j z_j^{p'} \leq 1\,, \text{then } \|\sum_j z_j \mathbf{1}_{T_j}\|_{p'} \lesssim \delta^{2(\frac{d}{p'}-(d-1))}\,.$$

Observe that $d/p' - (d-1) \nearrow 0$ as $p' \searrow d/(d-1)$. Thus, for any $\varepsilon > 0$, we have

$$\text{if } \delta^{d-1} \sum_j z_j^{p'} \leq 1\,, \text{then } \|\sum_j z_j \mathbf{1}_{T_j}\|_{p'} \lesssim \delta^{-\varepsilon}$$

if p' is close enough to $d/(d-1)$. So, by Lemma 15.6, this implies for any $\varepsilon > 0$

$$\|f_\delta^*\|_{L^p(\mathbb{S}^{d-1})} \lesssim_\varepsilon \delta^{-\varepsilon} \|f\|_{L^p(\mathbb{R}^d)}\,,$$

provided $p < d$ is close enough to d. Interpolating this with the trivial L^∞ bound yields the claimed estimate. $\square$

15.2. **Relation to the Kakeya set conjecture.** One consequence of the Kakeya maximal conjecture is the following statement concerning Besicovitch sets. Recall that such sets settle the d-dimensional Kakeya needle problem, i.e., they contain a unit line segment in every direction. Besicovitch's construction shows that such sets can have measure zero. However, it is not clear what their dimension is.

Conjecture 15.8 (Kakeya set conjecture)**.** *All Besicovitch sets have Hausdorff dimension and Minkowski dimension equal to d.*

Let us very briefly recall the definition of Hausdorff dimension (which is a bit tricky) and Minkowski dimension, at least for compact sets. See Appendix E for more details.

Definition 15.9 (Minkowski dimension)**.** Let E be a compact subset of $\mathbb{R}^d$. The set E is said to have Minkowski dimension n if

$$\lim_{\delta\to 0} \log_\delta |E_\delta| = d - n$$

where E_δ is the δ-neighborhood of E.

There are in fact two refined definitions.

Definition 15.10 (Upper Minkowski dimension)**.** The *upper Minkowski dimension* (or box packing dimension) $\overline{\dim}(E)$ of a set $E \subseteq \mathbb{R}^d$ is defined as the infimum over all exponents n such that for any $0 < \delta \ll 1$, the set E can be covered by $\mathcal{O}(\delta^{-n})$ balls of radius δ.

Definition 15.11 (Lower Minkowski dimension)**.** The *lower Minkowski dimension* $\underline{\dim}(E)$ is the infimum of all exponents n such that there exists arbitrarily small $0 < \delta \ll 1$ for which the set E can be covered by $\mathcal{O}(\delta^{-n})$ balls of radius δ.

Definition 15.12 (Hausdorff dimension)**.** The *Hausdorff dimension* $\dim_H(E)$ is defined as the infimum of all exponents n such that for any $0 < \delta \ll 1$, the set E can be covered by a countable collection of balls $B(x_i, r_i)$ of radius $r_i \le \delta$ such that $\sum_i r_i^n \lesssim 1$.

Clearly, $\dim_H(E) \le \underline{\dim}(E) \le \overline{\dim}(E)$, i.e., the Minkowski forms of the Kakeya conjecture are easier. For an introduction to Hausdorff measures, we refer to Appendix E and the references contained therein.

Proposition 15.13. *(1) The Kakeya maximal function conjecture implies the Kakeya set conjecture. More precisely, if (for $0 < \delta \ll 1$) it holds that*

$$\forall \varepsilon > 0 \; \exists C_\varepsilon : \; \|f_\delta^*\|_{L^p(\mathbb{S}^{d-1})} \le C_\varepsilon \delta^{-\varepsilon} \|f\|_{L^p(\mathbb{R}^d)} \tag{15.13}$$

for some $p < \infty$, then Besicovitch sets in $\mathbb{R}^d$ have Hausdorff dimension d.

(2) More generally, if

$$|\{\omega \in S^{d-1} : (\mathbf{1}_E)_\delta^*(\omega) > \alpha\}| \lesssim_\varepsilon \delta^{-d+p-\varepsilon} \alpha^{-p} |E| \tag{15.14}$$

holds for all $\varepsilon > 0$ and $0 < \delta < 1$, and $E \subseteq \mathbb{R}^d$ is a Borel set having the property that for each $\omega \in S^{d-1}$ there is a unit segment γ_ω in direction ω for which $|\gamma_\omega \cap E|_{\mathbb{R}} > 0$, then $\dim_H E \ge p$.

Remark 15.14. (1) The inequality

$$|E_\delta| \ge C_\varepsilon^{-1} \delta^\varepsilon \tag{15.15}$$

for any Kakeya set E (and its δ-neighborhood E_δ) follows immediately from (15.13) by the same argument that was used in (2) in Remark 15.3. Formula (15.15) says that Besicovitch sets in $\mathbb{R}^d$ have lower Minkowski dimension d.

(2) Note that (15.14) is a weaker version of (15.11) with $q = p$ and f being of the form $\mathbf{1}_E$.

Proof of Proposition 15.13. See Wolff [200, Proposition 10.2] for the first and Sogge [161, Proposition 9.1.5] for the second part.

(1) Let E be a Besicovitch set. Fom Remark E.2 (i.e., $H_\alpha(E) = 0$ for any $\alpha > d$) and the definition of the Hausdorff measure (Lemma E.3), it suffices to show that for a given covering of E by balls $B_j := B_{x_j}(r_j)$ with, say, $r_j \leq 1/100$, we have $\sum r_j^\alpha \gtrsim 1$ for any $\alpha < d$. For this let

$$J_k := \{j : \ 2^{-k} \leq r_j \leq 2^{-(k-1)}\}$$

and denote by I_ω any unit line segment oriented in the direction $\omega \in \mathbb{S}^{-1}$ which is contained in our Kakeya set E. Let further

$$S_k := \left\{ \omega \in \mathbb{S}^{d-1} : \ |I_\omega \cap \bigcup_{j\in J_k} B_j| \geq \frac{1}{100k^2} \right\} .$$

Since

$$\sum_k (100k^2)^{-1} < 1 \quad \text{and} \quad \sum_k |I_\omega \cap \bigcup_{j\in J_k} B_j| \geq |I_\omega| = 1 ,$$

we see $\bigcup_k S_k = \mathbb{S}^{d-1}$. (If not, we could find some $\omega_0 \notin S_k$ for every $k = 1, 2, ...$ meaning that $|I_{\omega_0} \cap \bigcup_{j\in J_k} B_j| \leq (100k^2)^{-1}$. But since $I_\omega \subseteq \bigcup_j B_j$ we must have $1 = |I_{\omega_0}| \leq \sum_{k\geq 1} |I_{\omega_0} \bigcup_j B_j| \leq \sum_{k\geq 1}(100k^2)^{-1} < 1$.) In particular, it is clear that $\sigma(S_k) \gtrsim 1$, where σ denotes the euclidean surface measure on $\mathbb{S}^{d-1}$.

Now, let

$$f = \mathbf{1}_{F_k} , \quad \text{where } F_k := \bigcup_{j\in J_k} B_{x_j}(10r_j) .$$

Then, for $\omega \in S_k$ we have (for a tube $T_\omega^\delta(a)$ of length 1 and thickness δ oriented along ω, and centered at $a \in \mathbb{R}^d$),

$$|T_\omega^{2^{-k}}(a_\omega) \cap F_k| \gtrsim \frac{|T_\omega^{2^{-k}}(a_\omega)|}{100k^2} ,$$

where a_ω denotes the midpoint of I_ω. Hence, after a short computation (see also the ensuing remark), we see

$$\|f_{2^{-k}}^*\|_p \gtrsim k^{-2}\sigma(S_k)^{1/p} . \tag{15.16}$$

On the other hand, (15.2) implies that

$$\|f_{2^{-k}}^*\|_p \leq C_\varepsilon 2^{k\varepsilon}\|f\|_p \leq C_\varepsilon 2^{k\varepsilon}(|J_k|2^{-(k-1)d})^{1/p} . \tag{15.17}$$

Comparing (15.16) and (15.17) therefore shows

$$\sigma(S_k) \lesssim 2^{kp\varepsilon - kd}k^{2p}|J_k| \lesssim 2^{-k(d-2p\varepsilon)}|J_k| .$$

Therefore,

$$\sum_j r_j^{d-2p\varepsilon} \geq \sum_k 2^{-k(d-2p\varepsilon)}|J_k| \gtrsim \sum_k \sigma(S_k) \gtrsim 1 ,$$

which was asserted at the beginning of the proof (for $0 < \alpha = p\varepsilon < d$ with $p < \infty$ and ε sufficiently small).

(2) See Sogge [161, Proposition 9.1.5]. We proceed similarly as in (1). One of our hypotheses is slightly weaker since we are not assuming that for each $\omega \in S^{d-1}$ we can find a unit segment γ_ω in this direction contained in E. However, since

$$S^{d-1} = \bigcup_{0<\alpha<1} \{\omega \in S^{d-1} : \exists \gamma_\omega \text{ with } |\gamma_\omega \cap E|_{\mathbb{R}} > \alpha\} ,$$

it follows that we can find $\alpha_0 \in (0,1)$ and $U \subseteq S^{d-1}$ so that $|U|_{S^{d-1}} > 0$ and that for each $\omega \in U$ there is a unit line segment in the direction of ω so that

$$|\gamma_\omega \cap E|_{\mathbb{R}} > \alpha_0 \,.$$

To use this, suppose that $E \subseteq B_{x_j}(r_j)$ is a covering by balls of radius $r_j \in (0,1/2)$ and let, as before, $J_k = \{j : 2^{-k} \leq r_j < 2^{-k+1}\}$ being the index set of those r_j satisfying $r_j \in [2^{-k}, 2^{-k+1})$. Then if now

$$U_k = \{\omega \in U : |\gamma_\omega \cap \bigcup_{j\in J_k} B_{x_j}(r_j)|_{\mathbb{R}} > \frac{\alpha_0}{\pi(1+k^2)}\} \,,$$

by the earlier argument where we showed $S^{d-1} = \bigcup_{k\geq 1} S_k$[15], we must have $U = \bigcup_{k=1}^\infty U_k$. If $D_k := \bigcup_{j\in J_k} B_{x_j}(2r_j)$, then we also get

$$(\mathbf{1}_{D_k})^*_{2^{-k}}(\omega) > \frac{\alpha_0}{2\pi(1+k^2)} \,, \quad \omega \in U_k \,.$$

Consequently, by (15.14) (with $E = D_k$, $\delta = 2^{-k}$, S^{d-1} replaced by U_k, and α replaced by $\alpha_0/[\pi(1+k^2)]$), we have

$$|U_k|_{S^{d-1}} \lesssim_{\varepsilon,\alpha_0} (1+k^2)^p 2^{k(n-p+\varepsilon/2)}|D_k| \lesssim_{\varepsilon,\alpha_0} 2^{-k(p-\varepsilon)}|J_k|$$

where $|J_k|$ denotes the cardinality of J_k and we used $|D_k| \lesssim |J_k| r_j^d \sim |J_k| 2^{-kd}$ as well as $(1+k^2)^p \lesssim 2^{k\varepsilon/2}$. Therefore, if $0 < \varepsilon < 1$, then summing this estimate over all j, we obtain

$$\sum_j r_j^{p-\varepsilon} \geq \sum_{k=1}^\infty \sum_{j\in J_k} 2^{-k(p-\varepsilon)}|J_k| \gtrsim_{\varepsilon,\alpha_0} \sum_{k=1}^\infty |U_k|_{S^{d-1}} \gtrsim_{\varepsilon,\alpha_0} |U|_{S^{d-1}} > 0 \,.$$

Hence, by Definition 15.9, see also Definition E.1 and Lemma E.3, we must have $\dim_H E \geq p$ as claimed (cf. [161, Lemma 9.1.3]). □

Remark 15.15. Let us quickly justify (15.16). Since $f = \mathbf{1}_{F_k}$ and $|T_\omega^{2^{-k}}(a) \cap F_k| \gtrsim k^{-2}|T_\omega^{2^{-k}}(a)|$, whenever $\omega \in S_k$, we have

$$f^*_{2^{-k}}(\omega) = \sup_{a\in\mathbb{R}^d} \frac{1}{|T_\omega^{2^{-k}}(a)|}\int_{T_\omega^{2^{-k}}(a)} \mathbf{1}_{F_k}(x) \gtrsim \sup_{a\in\mathbb{R}^d} \frac{|T_\omega^{2^{-k}}(a)|}{|T_\omega^{2^{-k}}(a)|} \cdot k^{-2}\mathbf{1}_{S_k}(\omega) \,.$$

Therefore,

$$\|f^*_{2^{-k}}\|_{L^p(\mathbb{S}^{d-1})} \gtrsim k^{-2}\|\mathbf{1}_{S_k}\|_{L^p(\mathbb{S}^{d-1})} = k^{-2}\sigma(S_k)^{1/p} \,,$$

where σ denotes the euclidean surface measure on $\mathbb{S}^{d-1}$.

Although appearing quite elementary, the Kakeya conjecture is a major open problem in geometric measure theory which is closely connected to many classical problems in Fourier analysis regarding estimation of oscillatory integrals. This is a consequence of Fefferman's solution of the disk multiplier problem [73] and work of Córdoba (e.g., [51]) and Bourgain (e.g., [15, 19, 27]). So far, the conjecture was only shown in $d = 2$ by an elegant argument of Córdoba and Fefferman. (See also Subsection 7.4 and [179, Lecture 5] for a proof using bilinear estimates.) Of course, the conjecture is also an immediate consequence of the two-dimensional restriction estimate (that we outlined in Subsection 7.4) and the square function estimate by Córdoba and Fefferman, see Appendix B.

For further information on this classical problem, we refer the reader to the excellent reviews by Wolff [198] and Katz and Tao [117]. Here, we shall content ourselves with the treatment of the problem in $d = 2$ and review the (direct!) proofs by Córdoba [51] (which is based on geometric

[15]See also the proof of [161, Theorem 9.1.4].

arguments) and Bourgain [15] (which uses Fourier analysis). Here, we follow again Wolff [200, Theorem 10.3].

Theorem 15.16. *If $d = 2$, then we have the bound*

$$\|f_\delta^*\|_{L^2(\mathbb{S}^1)} \leq C(\log(1/\delta))^{1/2}\|f\|_{L^2(\mathbb{R}^2)}\,.$$

Proof of Theorem 15.16 due to Bourgain. Without loss of generality, we can assume $f \geq 0$. Introducing

$$\rho_\delta^\omega(x) := \frac{1}{2\delta}\mathbf{1}_{T_\omega^\delta(0)}\,,$$

we see that the maximal function can be written as

$$f_\delta^*(\omega) = \sup_{a\in\mathbb{R}^2}(f * \rho_\delta^\omega)(a)\,.$$

Now let us find a pointwise upper bound on this function. To this end, we introduce $0 \leq \varphi \in \mathcal{S}(\mathbb{R})$ such that $\hat{\varphi}$ is compactly supported and $\varphi(x_1) \geq 1$ for $|x_1| \leq 1$. Let us further define

$$\begin{aligned}\psi &: \mathbb{R}^2 \to \mathbb{R}\\ x &\mapsto \varphi(x_1)\cdot\frac{1}{2\delta}\varphi(x_2/\delta)\,,\end{aligned}$$

i.e., a smoothed out characteristic function of a $\delta \times 1$ tube oriented along the e_1-axis. Note that $\psi(x) \geq \rho_\delta^{e_1}(x)$ and therefore $f_\delta^*(e_1) \leq \sup_{a\in\mathbb{R}^2}(f * \psi)(a)$. Thus, if we similarly define $\psi_\omega := \psi \circ p_\omega$ for some rotation $p_\omega \in \mathrm{SO}(2)$, we obtain similarly $\psi_\omega(x) \geq \rho_\delta^\omega(x)$. Using this bound and Cauchy–Schwarz, we can therefore estimate

$$|f_\delta^*(\omega)|^2 \leq |\sup_{a\in\mathbb{R}^2}(f*\psi_\omega)(a)|^2 \leq \|\hat{f}\cdot\widehat{\psi_\omega}\|_1^2 \leq \left(\int_{\mathbb{R}^2}|\hat{f}(\xi)|^2|\widehat{\psi_\omega}(\xi)| < \xi > \,d\xi\right)\cdot\left(\int_{\mathbb{R}^2}\frac{|\widehat{\psi_\omega}(\xi)|}{< \xi >}\,d\xi\right).$$

Now, since $\widehat{\psi_\omega} = \hat{\psi}\circ p_\omega$, we know that $\widehat{\psi_\omega}$ is supported on the dual rectangle R_ω (oriented along ω) with dimensions $|\xi_1| \sim 1$ and $|\xi_2| \sim \delta^{-1}$. Combining this with $|\hat{\psi}| \lesssim 1$, we obtain

$$\int_{\mathbb{R}^2}\frac{|\widehat{\psi_\omega}(\xi)|}{< \xi >}\,d\xi \lesssim \int_{R_\omega}\frac{d\xi}{< \xi >} \sim \int_1^{1/\delta} t^{-1}\,dt = \log(1/\delta)\,.$$

(If we considered the d-dimensional problem, we would get a factor $\delta^{-(d-2)}$ instead.) Now putting all estimates together, we obtain

$$\begin{aligned}\|f_\delta^*\|_{L^2(\mathbb{S}^1)}^2 &\lesssim \log\frac{1}{\delta}\int_{\mathbb{S}^1}d\omega\int_{\mathbb{R}^2}|\widehat{\psi_\omega}(\xi)||\hat{f}(\xi)|^2 < \xi > \,d\xi\\ &= \log\frac{1}{\delta}\int_{\mathbb{R}^2}|\hat{f}(\xi)|^2 < \xi > \left(\int_{\mathbb{S}^1}|\widehat{\psi_\omega}(\xi)|\,d\omega\right)d\xi\\ &\lesssim \log\frac{1}{\delta}\int_{\mathbb{R}^2}|\hat{f}(\xi)|^2\,d\xi = \log\frac{1}{\delta}\|f\|_2^2\,.\end{aligned}$$

To get from the second to the last line, we used that, for fixed $\xi \in \mathbb{R}^2$, the set of $\omega \in \mathbb{S}^{d-1}$ where $\widehat{\psi_\omega}(\xi) \neq 0$ holds, has measure $\lesssim < \xi >^{-1}$, see also the ensuing remark. This concludes the proof. □

Remarks 15.17. (1) As we remarked after the estimate of $\int_{\mathbb{R}^2}\frac{|\widehat{\psi_\omega}(\xi)|}{<\xi>}\,d\xi$, the above arguments show

$$\|f_\delta^*\|_{L^2(\mathbb{S}^{d-1})} \lesssim \delta^{-(d-2)/2}\|f\|_{L^2(\mathbb{R}^d)} \qquad (15.18)$$

in $d \geq 3$ dimensions, which is the best possible L^2 bound.

(2) Let us elaborate a bit more on the estimate

$$|\{\omega \in \mathbb{S}^{d-1} : \ |\widehat{\psi_\omega}(\xi)| > 0\}| \lesssim < \xi >^{-1} \tag{15.19}$$

for given (fixed) $\xi \in \mathbb{R}^d$. Recall that $\widehat{\psi_\omega}$ was a smoothed out (and compactly supported!) indicator function of a $\delta^{-1} \times \cdots \delta^{-1} \times 1$ rectangle, oriented in the direction ω and centered at the origin. Thus, to prove (15.19) we can pretend that $\widehat{\psi_\omega}$ is actually a smoothed indicator function of a thickened hyperplane with thickness $\mathcal{O}(1)$, say, e.g., 10. Moreover, by an elementary geometrical observation, it suffices to consider only the case $d = 2$. Next, by the underlying rotational symmetry, it suffices to consider only $\mathbb{R}^2 \ni \xi = (\xi_1, 0)$. Now suppose first that $|\xi| = \mathcal{O}(1)$, say $|\xi| \leq 10000$. Then, the left side of (15.19) is trivially bounded by $|\mathbb{S}^{d-1}|$ and so we are done in this case. In conclusion, we are left with estimating the left side of (15.19) in $d = 2$ when $\xi = |\xi|\hat{e}_1$ with $|\xi| \gg 1$ (say $|\xi| \geq 10000$), and $\widehat{\psi_\omega}(\xi)$ is replaced by the indicator function of an infinitely elongated tube of thickness 10, oriented along ω and centered at the origin. Let us for simplicity also assume that the tube is shifted in negative e_2-direction such that the upper border coincides with the e_1-axis. Then, as we start rotating the tube in positive direction with the rotation center being $(0, -1/2)$, there will be a rotation angle φ where the lower border of the tube touches ξ; that's precisely the angle, we are interested in since

$$|\{\omega \in \mathbb{S}^1 : \ |\widehat{\psi_\omega}(\xi)| > 0\}| \lesssim \int_{\mathbb{S}^1} \mathbf{1}_{T_\omega}(\xi) \sim \int_0^\varphi d\varphi' = \varphi \,.$$

But by elementary trigonometry, this angle is given by $\sin(\varphi/2) = 1/(2|\xi|)$. Since $|\xi| \gg 1$, we may approximate $\sin\varphi \sim \varphi$ which shows the claim, see also the following figure.

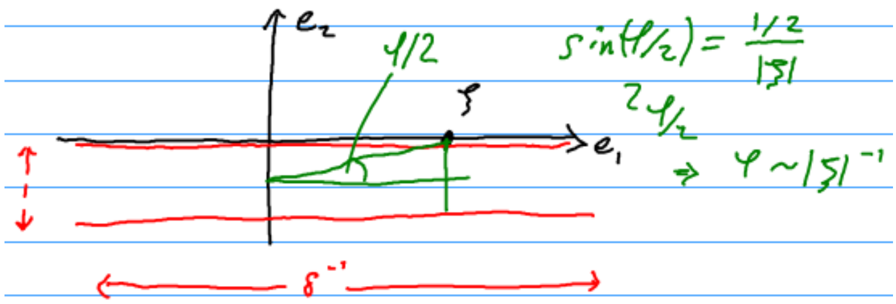

Figure 4

As opposed to Bourgain's proof, Córdoba elegantly exploited a simple geometric fact. Apart from a technical issue involving small angles, the main point is that two lines intersect in at most a point, whereas every two thin rectangles intersect in a small parallelogram.

Proof of Theorem 15.16 due to Córdoba. In view of the auxiliary Lemma 15.6, it suffices to prove that for any subsequence $\{y_k\}$ with $\delta \sum y_k^2 = 1$ and any maximal δ-separated subset $\{e_k\}$ of $\mathbb{S}^1$, we have

$$\left\| \sum_k y_k \mathbf{1}_{T_{e_k}^\delta(a_k)} \right\|_2 \lesssim (\log(1/\delta))^{1/2} \,. \tag{15.20}$$

The relevant geometric fact is

$$|T_{e_k}^\delta(a) \cap T_{e_\ell}^\delta(b)| \lesssim \frac{\delta^2}{|e_k - e_\ell| + \delta} \,, \tag{15.21}$$

which is proven in the remark below. Using this, we can estimate the left side of (15.20)

$$\begin{aligned}\Big\|\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\Big\|_2^2 &= \sum_{k,\ell} y_k y_\ell |T^\delta_{e_k}(a_k)\cap T^\delta_\ell(a_\ell)| \lesssim \sum_{k,\ell} y_k y_\ell \frac{\delta^2}{|e_k-e_\ell|+\delta}\\ &= \sum_{k,\ell}\sqrt{\delta}y_k\cdot\sqrt{\delta}y_\ell\frac{\delta}{|e_k-e_\ell|+\delta} \le \|\sqrt{\delta}y_k\|_{\ell^2_k}\,\|\sum_\ell K_{k,\ell}\sqrt{\delta}y_\ell\|_{\ell^2_k}\,.\end{aligned}$$

Here we abbreviated

$$K_{k,\ell} := \frac{\delta}{|e_k-e_\ell|+\delta}$$

and denoted by ℓ^2_k the usual ℓ^2 space where the summation is with respect to k. Now recall that the set of $\{e_k\}$ is maximal δ-separated. Thus, for fixed k, there are at most δ^{-1} summands in the ℓ-summation. Moreover, since the angle between e_k and e_ℓ is given by $\delta|k-\ell|$, we have

$$|e_k-e_\ell| = \sqrt{2}\sqrt{1-\cos(\delta|k-\ell|)} \ge \sqrt{\frac{2}{100}}\delta|k-\ell| \quad \text{for } |k-\ell|<\frac{1}{\delta}\,.$$

Therefore, we can estimate

$$\sup_k\sum_\ell \frac{\delta}{|e_k-e_\ell|+\delta} \lesssim \sum_{\ell\le 1/\delta}\frac{\delta}{\delta\ell+\delta} \sim \log\frac{1}{\delta}\,.$$

Thus, we can apply Schur's test (Lemma 4.18) to the kernel $K_{k,\ell}$ (which is symmetric in k and ℓ) which allows us to estimate the left side of (15.20) further by

$$\Big\|\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\Big\|_2^2 \lesssim \log\frac{1}{\delta}\sum_k(\sqrt{\delta}y_k)^2 \lesssim \log\frac{1}{\delta}\,.$$

In view of the hypothesis $\delta\sum y_k^2 = 1$ (recall Lemma 15.6), this concludes the proof. □

Remark 15.18. Let us shortly elaborate on the measure of the intersection of two thin rectangles in (15.21) which can be reformulated as

$$|T^\delta_{e_k}(a)\cap T^\delta_{e_\ell}(b)| \lesssim \min\left\{\delta, \frac{\delta^2}{|e_k-e_\ell|}\right\}\,.$$

Clearly, it suffices to consider $a=b$, $e_\ell = e_1 \equiv (1,0)$, and that the angle θ between $e_k \equiv (\cos\theta,\sin\theta)$ and e_1 is at most $\pi/2$. Since $|T^\delta_{e_k}(a)|\le\delta$, the first bound is trivial. Now suppose $|e_k-e_1| = \sqrt{2}\sqrt{1-\cos\theta}\ge\delta$, which is (since $\cos\theta\le 1-\theta^2/4$) satisfied if $\theta>\sqrt{2}\delta$. In this case, $\sin\theta>\theta/2>\delta/2$ and we have $\delta/(2\sin\theta)<1$. Using $\cos\theta\ge 1-\theta^2$ and the formula for the surface area of a parallelogram, we finally obtain

$$|T^\delta_{e_1}\cap T^\delta_{e_k}| = 4\frac{\delta}{2}\cdot\frac{\delta}{2\sin\theta} \le \frac{2\delta^2}{\theta} \le \frac{2\sqrt{2}\delta^2}{\sqrt{2}\sqrt{1-\cos\theta}} = \frac{2\sqrt{2}\delta^2}{|e_k-e_1|} \tag{15.22}$$

what had to be proven.

15.3. **Universal bounds for the Kakeya maximal operator.** We follow Sogge [161, Section 9.2]. The reason why these bounds are called "universal" is that they are indeed optimal in curved spaces. Recall that Wolff [197] found improved bounds in the euclidean setting. (He got the following theorem for $p\le(d+2)/2$ instead of $p\le(d+1)/2$.) See also Subsection 15.4.

The main goal of this subsection is to prove non-trivial bounds for the Kakeya maximal function in higher dimensions using Bourgain's bush method. An improved bound is due to Wolff [197] (see also [161, Theorem 9.4.1]) who could at least treat $p=(d+2)/2$ giving the critical exponent in $d=2$ in the following theorem.

Theorem 15.19. *Let $d \geq 3$. Given $\varepsilon > 0$ and $0 < \delta < 1/2$, we have*

$$\|f_\delta^*\|_{L^q(S^{d-1})} \lesssim_\varepsilon \delta^{-\frac{d}{p}+1-\varepsilon} \|f\|_{L^p(\mathbb{R}^d)} \tag{15.23}$$

whenever $1 \leq p \leq (d+1)/2$ and $q = (d-1)p'$.

Observe (or recall) that the trivial case $p = 1$, i.e.,

$$\|f_\delta^*\|_{L^\infty(S^{d-1})} \leq \delta^{-d+1} \|f\|_{L^1(\mathbb{R}^d)}$$

If $p = (d+1)/2$, then $q = d+1$. To prove the other estimates, recall

Theorem 15.20 (General Marcinkiewicz interpolation)**.** *Suppose T is a subadditive operator of restricted weak types (p_j, q_j) with $p_0 < p_1$ and $q_0 \neq q_1$, i.e.,*

$$\|T\mathbf{1}_E\|_{L^{q_j,\infty}} \lesssim \|\mathbf{1}_E\|_{L^{p_j,1}} \sim |E|^{1/p_j} .$$

Then one has the estimate

$$\|Tf\|_{L^{q_\theta,r}} \lesssim \|f\|_{L^{p_\theta,r}} ,$$

for all $1 \leq r \leq \infty$, $\theta \in (0,1)$ with $q_\theta > 1$. If additionally $q_\theta \geq p_\theta$ and $r = q_\theta$, then

$$\|Tf\|_{L^{q_\theta}} \lesssim \|f\|_{L^{p_\theta}} .$$

Proof. See Stein–Weiss [169, Chapter V, Theorem 3.15] or Theorem 1.3.4 in harmonic analysis notes of summer term 2020. □

Proof of Theorem 15.19. Since the case $p = 1$ in (15.23) is trivial, it suffices, by the above Marcinkiewicz interpolation, to show the corresponding restricted weak-type $((d+1)/2, 1) \to (d+1, \infty)$ bound

$$|\{\omega \in S^{d-1} : (\mathbf{1}_E)_\delta^*(\omega) > \alpha\}|^{\frac{1}{d+1}} \lesssim \alpha^{-1} \delta^{-\frac{d-1}{d+1}} |E|^{\frac{2}{d+1}} ,$$

that is

$$|\{\omega \in S^{d-1} : (\mathbf{1}_E)_\delta^*(\omega) > \alpha\}| \lesssim \alpha^{-(d+1)} \delta^{-(d-1)} |E|^2 . \tag{15.24}$$

For a constant $A > 1$ to be fixed later, we set

$$\Omega_\alpha := \{\omega \in S^{d-1} : (\mathbf{1}_E)_\delta^*(\omega) > A\alpha\} .$$

Then we would have (15.24) once we prove

$$|\Omega_\alpha|_{S^{d-1}} \lesssim \alpha^{-(d+1)} \delta^{-(d-1)} |E|^2 , \quad \alpha > 0 , \quad 0 < \delta < 1/2 . \tag{15.25}$$

At the end of the proof we shall see that the case where $\alpha \lesssim \delta$ is trivial. Thus, let us, for the moment at least, also assume that $\alpha > A\delta$.

Now choose a maximal $(A\delta/\alpha)$-separated subset $\{\omega_j\}_{j=1}^M \equiv \mathcal{I}$ in Ω_α. Then it follows that

$$|\Omega_\alpha|_{S^{d-1}} \lesssim_d (A\delta/\alpha)^{d-1} M . \tag{15.26}$$

Thus, to get the desired bound on $|\Omega_\alpha|$, we need good bounds on the number M.

If $\omega_j \in \mathcal{I}$, then, by definition of Ω_α, we have

$$|E \cap T_{\omega_j}| > A\alpha |T_{\omega_j}| \tag{15.27}$$

and so, by summing over j (and recalling $|T_\omega| = \delta^{d-1}$), we have

$$\sum_{j=1}^M |E \cap T_{\omega_j}| \geq c_0(A,d) M \alpha \delta^{d-1}$$

for some $c_0 = c_0(A, d)$. Thus,

$$\frac{1}{|E|}\int_E \sum_{j=1}^{M} \mathbf{1}_{T_{\omega_j}} \geq \frac{c_0(A,d)M\alpha\delta^{d-1}}{|E|}\,.$$

Since there must be a point $a \in E$ where the non-negative function $\sum_{j=1}^{M} \mathbf{1}_{T_{\omega_j}}$ equals or exceeds its average over E[16], i.e.,

$$\frac{1}{|E|}\int_E \sum_{j=1}^{M} \mathbf{1}_{T_{\omega_j}} \leq \sum_{j=1}^{M} \mathbf{1}_{T_{\omega_j}}(a)\,,$$

we obtain

$$\sum_{j=1}^{M} \mathbf{1}_{T_{\omega_j}}(a) \geq \frac{c_0(A,d)M\alpha\delta^{d-1}}{|E|}\,, \quad \text{some } a \in E\,.$$

Put differently, by the pigeonhole principle, this point $a \in E$ must belong to at least $\mathbb{N} \ni N \leq M$ tubes $\{T_{\omega_j}\}_{j=1}^{M}$ such that

$$N \geq \frac{c_0(A,d)M\alpha\delta^{d-1}}{|E|}\,. \tag{15.28}$$

Let us collect these tubes from the original collection into the "bush" centered at $a \in \mathbb{R}^d$ (see Figure 5)

$$\{T_{\omega_{j_k}}\}_{k=1}^{N}\,.$$

Now, since the points $\omega_j \in S^{d-1}$ are $(A\delta/\alpha)$-separated and since, for the moment, we are assuming $\alpha > A\delta$, we conclude that if A is large enough, we must have[17]

$$\left(T_{\omega_{j_k}} \cap T_{\omega_{j_\ell}}\right) \setminus B_a(\alpha) = \emptyset\,, \quad \text{if } k \neq \ell\,.$$

Therefore, the "tips" of the branches of the bush about $a \in \mathbb{R}^d$, denoted by

$$\tau_{j_k} := T_{\omega_{j_k}} \setminus B_a(\alpha)\,, \quad 1 \leq k \leq N\,,$$

[16]This point $a \in \mathbb{R}^d$ is a point where a preferably large number of tubes T_{ω_j} intersect themselves as well as the set E. The latter property is not that important for the moment; the former point means that there is a subcollection of the intersecting T_{ω_j} that form a "Bourgain bush", see Figure 5.

[17]This is a simple consequence of the geometrical fact, that if ℓ_1 and ℓ_2 are two lines crossing each other at the origin with angle $\theta \in (0, \pi/2]$, then $\mathrm{dist}(\ell_1 \cap rS^{d-1}, \ell_2 \cap rS^{d-1}) \sim r\theta$. In our case, $r = \alpha$ and $\theta = A\delta/\alpha$, so the distance on the sphere is roughly $A\delta > 0$.

are disjoint as depicted in Figure 5.

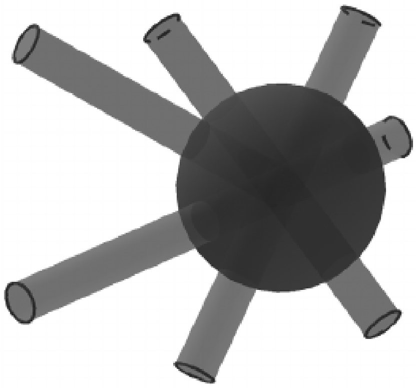

Figure 5. A Bourgain bush

Since

$$|T_{\omega_j} \cap B_a(\alpha)| \le C_0\alpha |T_{\omega_j}|$$

for a uniform constant C_0, we conclude that if $A \ge 2C_0$ as well,

$$|\tau_{j_k} \cap E| \ge A\alpha |T_{\omega_{j_k}}| - C_0\alpha |T_{\omega_{j_k}}| \ge A\alpha |T_{\omega_{j_k}}|/2\,, \quad 1 \le k \le N$$

by (15.27). If we use this, the disjointness of the tips of the branches, and (15.28), we conclude

$$|E| \ge \sum_{k=1}^{N} |\tau_{j_k} \cap E| \ge c_d A\alpha\delta^{d-1} N \ge c_d' M\alpha^2\delta^{2(d-1)}/|E|\,,$$

or equivalently,

$$M \le C\alpha^{-2}\delta^{-2(d-1)}|E|^2\,. \tag{15.29}$$

If we plug this into (15.26), we obtain the desired bound

$$|\Omega_\alpha|_{S^{d-1}} \lesssim \alpha^{-(d+1)}\delta^{-(d-1)}|E|^2\,, \quad \alpha > 0\,, \quad 0 < \delta < 1/2 \tag{15.30}$$

stated ad the beginning of the proof in (15.25).

We are left with the case $\alpha < A\delta$ which is a lot easier. For assuming that $\Omega_\alpha \ne \emptyset$, we just use the fact that we can find a single tube T_ω so that

$$\alpha\delta^{d-1} \sim \alpha|T_\omega| \le |E \cap T_\omega| \le |E|\,.$$

Thus, there must be $c_0 > 0$ such that

$$c_0 \le \alpha^{-2}\delta^{-2(d-1)}|E|^2\,.$$

But since the right side is dominated by the right side of (15.30) if $\alpha < A\delta$, we conclude (15.30) must be valid in this case as the left side of (15.30) is at most $|S^{d-1}|$. ☐

15.4. **Wolff's bounds for the Kakeya maximal operator in higher dimensions.** We present the proof of Wolff's [197] bound for the Kakeya maximal function in higher dimensions and follow Sogge [161, Section 9.4]. When $d = 2$, the following theorem is optimal as we have seen earlier.

Theorem 15.21. *Let $d \ge 3$ and f_δ^* denote the Kakeya maximal function defined in* (15.5)*. Then given $\varepsilon > 0$ and $0 < \delta < 1/2$, we have*

$$\|f_\delta^*\|_{L^q(S^{d-1})} \lesssim_\varepsilon \delta^{-d/p+1-\varepsilon}\|f\|_{L^p(\mathbb{R}^d)} \tag{15.31}$$

whenever $1 \le p \le (d+2)/2$ *and* $q = (d-1)p'$*. In particular (by Proposition 15.13), we have that* $\dim_H E \ge (d+2)/2$ *for Besicovitch sets* $E \subseteq \mathbb{R}^d$*.*

15.5. **How can Kakeya help in proving the restriction conjecture?** In the first subsection we saw that the restriction conjecture implies the Kakeya maximal conjecture. Bourgain [15, Section 6] partially reversed this and obtained a restriction theorem beyond Tomas–Stein by using a Kakeya set estimate that is stronger than the L^2 bound stated in Wolff [200, Formula (151)], i.e.,

$$\|\sum_k y_k \mathbf{1}_{T^\delta_{e_k}(a_k)}\|_2^2 \lesssim \log\frac{1}{\delta}\ \sum_k (\sqrt{\delta} y_k)^2 \lesssim \log\frac{1}{\delta}\,,$$

used in the proof of the L^2 bound (15.18) (via Lemma 15.6). It is not known whether (either version of) the Kakeya conjecture implies the full restriction conjecture. Anyway, we have

Theorem 15.22 (Bourgain [15])**.** *Suppose that we have an estimate*

$$\|\sum_j \mathbf{1}_{T^\delta_{e_j}}\|_{q'} \le C_\varepsilon \delta^{-(\frac{d}{q}-1+\varepsilon)} \tag{15.32}$$

for any given $\varepsilon > 0$ *and for some fixed* $q > 2$*. Then*

$$\|\widehat{fd\sigma}\|_p \lesssim_p \|f\|_{L^\infty(\mathbb{S}^{d-1})} \tag{15.33}$$

for some $p < 2(d+1)/(d-1)$*.*

Remarks 15.23. (1) The geometrical statement corresponding to (15.32) is that Kakeya sets in $\mathbb{R}^d$ have Hausdorff dimension at least q, recall the second assertion in Proposition 15.13 and the ensuing remark.

(2) Note that (15.12), which stated $\|\widehat{fd\sigma}\|_p \lesssim_p \|f\|_{L^p(\mathbb{S}^{d-1})}$ for $p > 2d/(d-1)$. is only ostensibly stronger than (15.33). In fact, these estimates are (formally at least) equivalent, see Bourgain [15].

We shall sketch the proof only for $d = 3$ and follow Wolff [200, Theorem 10.6]. Recall that in this case, we already know the bounds

$$\|\widehat{fd\sigma}\|_{L^4(\mathbb{R}^3)} \lesssim \|f\|_{L^2(\mathbb{S}^2)}$$

from the Tomas–Stein theorem, and

$$\|\widehat{fd\sigma}\|_{L^2(B_0(R))} \lesssim R^{1/2}\|f\|_{L^2(\mathbb{S}^2)}$$

from Theorem 4.17 with $\alpha = d-1 = 2$. Interpolating between these two estimates yields a family of estimates

$$\|\widehat{fd\sigma}\|_{L^p(B_0(R))} \lesssim R^{\frac{2}{p}-\frac{1}{2}}\|f\|_{L^2(\mathbb{S}^2)}\,, \quad \text{for } 2 \le p \le 4\,. \tag{15.34}$$

In the following argument we show that the exponent of R can be lowered by an ε if the L^2 norm on the right side is replaced by the L^∞ norm.

Proposition 15.24. *Let* $d = 2$*,* $2 < p < 4$*, and assume* (15.32) *holds for some* $q > 2$*. Then, for all* $\varepsilon > 0$*, we have*

$$\|\widehat{fd\sigma}\|_{L^p(B_0(R))} \lesssim_\varepsilon R^{\alpha(p)}\|f\|_{L^\infty(\mathbb{S}^{d-1})}\,, \quad \alpha(p) < \frac{2}{p} - \frac{1}{2}\,. \tag{15.35}$$

Clearly, this implies (15.33) for all p such that $\alpha(p) \le 0$, i.e., in particular, there are $p < 4$ for which (15.33) holds.

Heuristic proof of the proposition. By homogeneity, we can assume $\|f\|_{L^\infty(\mathbb{S}^2)} = 1$. Let $\delta = R^{-1}$ and cover $\mathbb{S}^2$ by the spherical caps

$$S_j = \{\omega \in \mathbb{S}^2 : \ |1 - \omega \cdot e_j| \le \delta\}\,,$$

where $\{e_j\}$ now forms a maximal $\delta^{1/2}$-separated subset of $\mathbb{S}^2$. Then we decompose

$$f = \sum_j f_j\,,$$

where each f_j is a Knapp example supported on S_j. Abbreviate $G = \widehat{fd\sigma}$ and $G_j = \widehat{f_j d\sigma}$ so that $G = \sum_j G_j$. By the uncertainty principle, the G_j are roughly constant on $\delta^{-1/2} \times \delta^{-1/2} \times \delta^{-1}$ tubes τ_j oriented along e_j and decaying rapidly away from them. For simplicity, let us assume in the following that G_j are in fact supported only on the τ_j [18].

Next, let us cover $B_0(R)$ with disjoint cubes Q of sidelength $\sqrt{R}$. For each fixed cube Q let $N(Q)$ denote the number of tubes τ_j that intersect Q. Note that $G|_Q = \sum_j G_j|_Q$, where we sum only over those j's for which τ_j intersects Q. Using this and the known restriction estimates (15.34), we can estimate $\|G\|_{L^p(Q)}$ for $2 \le p \le 4$ by

$$\begin{aligned}\|\widehat{fd\sigma}\|_{L^p(Q)} = \|G\|_{L^p(Q)} &\lesssim R^{\frac12\left(\frac2p - \frac12\right)} \Bigg\| \sum_{j:\tau_j \cap Q \neq \emptyset} f_j \Bigg\|_{L^2(\mathbb{S}^2)} \lesssim R^{\frac12\left(\frac2p - \frac12\right)} \left(N(Q)\|f\|\right)^{1/2} \\ &\sim \delta^{\frac34 - \frac1p} N(Q)^{1/2}\end{aligned}$$

where we used that the f_j are essentially disjointly supported and $\|f_j\|_{L^2(\mathbb{S}^2)} \sim |S_j|^{1/2} \sim \delta^{(d-1)/4} = \delta^{1/2}$. Summing over all Q then yields

$$\|\widehat{fd\sigma}\|^p_{L^p(B_0(R))} \lesssim \delta^{\frac{3p}{4} - 1} \sum_Q N(Q)^{p/2} \sim \delta^{\frac{3p}{4} + \frac12} \|\sum_j \mathbf{1}_{\tau_j}\|^{p/2}_{p/2} \tag{15.36}$$

where we used

$$\|\sum_j \mathbf{1}_{\tau_j}\|^{p/2}_{p/2} = \sum_Q N(Q)^{p/2}|Q| = \delta^{-3/2} \sum_Q N(Q)^{p/2}\,.$$

Now let $p = 2q'$, where q' is the exponent in (15.32), and assume that p is sufficiently close to 4 (and interpolate between (15.32) and (15.18), i.e., $\|f^*_\delta\|_{L^2(\mathbb{S}^2)} \lesssim \log(1/\delta)^{1/2}\|f\|_2$ if necessary). For any $\varepsilon > 0$, we have from the hypothesized, strengthened Kakeya set estimate (15.32),

$$\|\sum_j \mathbf{1}_{T^{\sqrt{\delta}}_{e_j}}\|_{q'} \le C_\varepsilon \delta^{-\frac12\left(\frac3q - 1 + \varepsilon\right)}\,.$$

Rescaling this inequality by δ^{-1} yields

$$\|\sum_j \mathbf{1}_{\tau_j}\|_{q'} \le C_\varepsilon \delta^{-\left(\frac3q - 1 + \varepsilon\right)} \cdot \delta^{-3/q'} = \delta^{-1-3/p-\varepsilon}\,.$$

Plugging this into (15.36) shows

$$\|\widehat{fd\sigma}\|_{L^p(B_0(R))} \lesssim \delta^{\frac14 - \frac1p - \varepsilon} = R^{\frac1p - \frac14 + \varepsilon}$$

and thereby the assertion since $1/p - 1/4 < 2/p - 1/2$ for $p < 4$. □

[18] It is precisely because of this assumption that our proof is merely heuristic. Clearly, the Fourier transform of a compactly supported measure cannot be compactly supported; the rigorous proof uses Schwartz decay of the G_j instead

In the next subsection, we will review Tao's surprising finding [181] that the Bochner–Riesz conjecture actually implies the restriction conjecture, thereby implying of course the Kakeya conjecture. In that context, we shall also review older work by Bourgain who directly proved the implication Bochner–Riesz ⇒ Kakeya. The latter is frequently used to construct counterexamples to L^p-boundedness of certain multipliers. The most prominent example being the failure of L^p-boundedness of the disk multiplier when $p \neq 2$ (Fefferman [73]).

16. Connection to the Bochner–Riesz conjecture

Historically, the first connection between – apparently purely geometrically involving considerations – the Kakeya conjecture and (Fourier) analysis arose in the 1970s. Considering the classical Fourier transform of a test function f in $\mathbb{R}^d$, one may ask whether the truncation

$$(S_R f)(x) := \int\limits_{|\xi|\leq R} \hat{f}(\xi) \mathrm{e}^{2\pi i x\cdot\xi}\, d\xi \tag{16.1}$$

converges as $R \to \infty$ to f in a certain sense, e.g., in L^p, or even pointwise almost everywhere. The above operator is usually referred to as the *ball multiplier* (*disk multiplier* in $d = 2$). Proofs of such assertions typically lie in proving the assumptions of the following two classical functional analytic results. Their proofs can be found, e.g., in Krantz [126, p. 27].

Lemma 16.1 (Functional analysis principle 1)**.** *Let X be a Banach space and $\mathcal{S}$ a dense subset. Let $T_R : X \to X$ be a sequence of linear operators (bounded on X) such that $T_R f \to Tf$ in X norm as $R \to \infty$ for test functions $f \in \mathcal{S}$ and some linear operator T that is also bounded on X. Then, in order to have $T_R f \to Tf$ in X norm for all functions in X (and not only test functions), it is a necessary and a sufficient condition to have the estimate*

$$\|T_R f\|_X \lesssim \|f\|_X \quad \text{for all sufficiently large } R \text{ and } f \in X\,.$$

Lemma 16.2 (Functional analysis principle 2)**.** *Let $1 \leq p < \infty$, $T_R : L^p \to L^p$ be a sequence of linear (L^p bounded) operators, and denote by*

$$(T^* f)(x) = \sup_R |(T_R f)(x)|$$

the maximal function *associated to T_R. Let $\mathcal{S} \subseteq L^p$ be a dense subset. Assume that*

1. *For each $s \in \mathcal{S}$, the limit $\lim_{R\to\infty}(T_R s)(x) \equiv (Ts)(x)$ exists in $\mathbb{C}$ for almost all $x \in \mathbb{R}^d$ and another L^p bounded operator T.*
2. *The associated maximal operator T^* has weak type (p, p), i.e. for each $\alpha > 0$,*

$$|\{x \in \mathbb{R}^d : (T^* f)(x) > \alpha\}| \lesssim \alpha^{-p}\|f\|_p \quad \text{for all } f \in L^p\,.$$

Then, for each $f \in L^p$, $\lim_{R\to\infty}(T_R f)(x)$ exists for almost all $x \in \mathbb{R}^d$.

Remark 16.3. The above two lemmas are by now standard tools to establish norm or pointwise almost everywhere convergence theorems. It is therefore natural to ask whether they are also strictly necessary. In particular, is it possible to have a convergence result $\lim_{R\to\infty} T_R f = Tf$ *without* being able to obtain uniform operator norm bound or a weak-type maximal inequality of the above forms?

In case of *norm* convergence, the answer is "no", thanks to the uniform boundedness principle, which among other things shows that norm convergence is only possible if one has the uniform bound

$$\|T_R f\|_X \lesssim \|f\|_X \quad \text{for all } f \in X\,,$$

see the proof of Lemma 16.1.

Returning to the pointwise almost everywhere convergence, the answer is in general "yes". Consider for instance the rank one operators

$$(T_n f)(x) := \int_{\mathbb{R}} \mathbf{1}_{[n,n+1]}(x-y) f(y)\, dy$$

from $L^1(\mathbb{R})$ to $L^1(\mathbb{R})$. It is clear that $\lim_{n\to\infty}(T_n f)(x) = 0$ almost everywhere for $f \in L^1(\mathbb{R})$ and that the operators T_n are uniformly bounded in L^1. However, the maximal function $T^* f$ does not belong to $L^{1,\infty}(\mathbb{R})$. One can modify this example in a number of ways to defeat almost any reasonable conjecture that something like the maximal weak-type estimate should be necessary for pointwise almost everywhere convergence. In spite of this, a remarkable observation of Stein [162], now known as *Stein's maximal principle*, asserts that the maximal weak-type inequality *is* necessary to prove pointwise almost everywhere convergence, if one is working on a compact group, the operators T_n are translation invariant, and the exponent p is at most 2.

Theorem 16.4 (Stein maximal principle). *Let G be a compact group, X be a homogeneous space of G with finite Haar measure μ, $1 \le p \le 2$, and $T_n : L^p(X) \to L^p(X)$ be a sequence of bounded linear operators commuting with translations such that $T_n f$ converges pointwise almost everywhere for each $f \in L^p(X)$. Then T^* has weak type (p,p).*

On the other hand, the theorem does fail for $p > 2$, and almost everywhere convergence results in L^p for $p > 2$ can be proven by other methods than weak (p,p) estimates. For instance, the convergence of Bochner–Riesz multipliers in $L^p(\mathbb{R}^d)$ for any d and for p in the range predicted by the Bochner–Riesz conjecture was verified by Carbery, Rubio de Francia, and Vega [39] (see Carbery [35] for $d = 2$ where, however, he proves a maximal weak-type inequality) despite the fact that the weak-type (p,p) estimate of even a *single* Bochner–Riesz multiplier, let alone the maximal function, has still not been completely verified in this range, especially for $1 < p < 2$, but see Tao [184] and Li and Wu [130] for maximal weak-type estimates in this range. (Carbery et al use weighted L^2 estimates for the maximal Bochner–Riesz operator, rather than L^p type estimates.) For $p \le 2$ though, Stein's principle (after localizing to a torus) does apply, and pointwise almost everywhere convergence of the Bochner–Riesz means is equivalent to the maximal weak-type (p,p) estimate.

Stein's principle is restricted to compact groups (such as the torus $(\mathbb{R}/\mathbb{Z})^d$ or the rotation group $SO(d)$) and their homogeneous spaces (such as the torus $(\mathbb{R}/\mathbb{Z})^d$ again, or the sphere $\mathbb{S}^{d-1}$), i.e., the principle fails in the non-compact setting (as in $\mathbb{R}$, as we have seen it before when dealing with $T_n f := f * \mathbf{1}_{[n,n+1]}$; the $T_n f$ converge pointwise almost everywhere to zero for every $f \in L^1(\mathbb{R})$, but the maximal function does not obey the weak-type $(1,1)$ estimate). However, in many applications on non-compact domains, the T_n are somewhat "localized" enough that one can transfer from a non-compact setting to a compact setting and then apply Stein's maximal principle. (For instance, Carleson's theorem [41] (see also Fefferman [74] for an alternative proof, Grafakos' book [96, Section 3.6.5] and `https://en.wikipedia.org/wiki/Carleson%27s_theorem` for references of expositions of Carleson's paper) on pointwise almost everywhere convergence of the partial Fourier series $\sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x}$ for $f \in L^2(\mathbb{R})$ is equivalent to Carleson's theorem on the circle $\mathbb{R}/\mathbb{Z}$ (due to the localization of the Dirichlet kernels) which is, due to Stein's principle, equivalent to a maximal weak-type $(2,2)$ estimate on the circle $\mathbb{R} \setminus \mathbb{Z}$. By a scaling argument in turn, this is equivalent to the analogous weak-type $(2,2)$ estimate on $\mathbb{R}$.)

See also Guzmán [61] for a systematic discussion of this and other maximal principles as well as `www.terrytao.wordpress.com/2011/05/12/steins-maximal-principle/` for more details. □

At this stage, it is also reasonable to remind the reader of the following sledge hammer whose proof can be found in Dunford and Schwartz [68] (Section XIII.6: Lemma 7 (p. 676), Theorem

8 (p. 678); Section XIII.8: Lemma 6 (p. 690) Theorem 7 (p. 693); Section XIII.9: Exercise 3 (p. 717)). The form which we shall use the theorem is as in [166, p. 48].

Lemma 16.5 (Hopf–Dunford–Schwartz ergodic theorem). *Let $\{T^t\}_{t\geq 0}$ be (measurable) semi-group of operators on $L^p(\mathbb{R}^d)$. Suppose that $\|T^t f\|_p \leq \|f\|_p$ for any $p \in [1,\infty]$. Then the maximal function*

$$(Mf)(x) = \sup_{s>0} \left(\frac{1}{s} \int_0^s |(T^t f)(x)|\, dt \right)$$

satisfies the inequalities

(1) $\|Mf\|_p \lesssim_p \|f\|_p$ *for all* $p \in (1,\infty]$;

(2) $|\{x \in \mathbb{R}^d : (Mf)(x) > \alpha\}| \lesssim \alpha^{-1}\|f\|_1$ *for each* $\alpha > 0$ *and* $f \in L^1$.

Proof of Lemma 16.1. Let $f \in X$ and suppose $\varepsilon > 0$. Then there exists an $s \in \mathcal{S}$ such that $\|f - s\| < \varepsilon$. Now select J so large that if $j,k \geq J$, then $\|T_j s - T_k s\| < \varepsilon$. For such j,k, we calculate

$$\begin{aligned}\|T_j f - T_k f\| &\leq \|T_j f - T_j s\| + \|T_j s - T_k s\| + \|T_k s - T_k f\| \\ &\leq \|T_j\|\, \|f - s\| + \varepsilon + \|T_k\|\, \|s - f\| \leq 3\varepsilon(1 + \sup_{\ell \geq J} \|T_\ell\|) \to 0 \quad \text{as } \varepsilon \to 0\,,\end{aligned}$$

i.e., $T_j f$ is Cauchy. Since X was supposed to be a Banach space, this establishes the result. The converse follows from the uniform boundedness principle, see, e.g., Rudin [148, p. 98] or Lieb and Loss [132, Theorem 2.12]. □

Proof of Lemma 16.2. The proof parallels that of Lemma 16.1 but is a bit more technical.

Let $f \in L^p$ and suppose that $\delta > 0$ is given. Then there is an $s \in \mathcal{S}$ such that $\|f - s\|_p^p < \delta$. For simplicity, we assume that both f and $T_j f$ are real-valued (the complex-valued case then follows from linearity). Fix $\varepsilon > 0$, independent of δ. Then

$$\begin{aligned}&|\{x : |\limsup_{j\to\infty}(T_j f)(x) - \liminf_{j\to\infty}(T_j f)(x)| > 3\varepsilon\}| \\ &\quad \leq |\{x : |\limsup_{j\to\infty}(T_j(f-s))(x)| > \varepsilon\}| + |\{x : |\limsup_{j\to\infty}(T_j s)(x) - \liminf_{j\to\infty}(T_j s)(x)| > \varepsilon\}| \\ &\qquad + |\{x : |\limsup_{j\to\infty}(T_j(s-f))(x)| > \varepsilon\}| \\ &\quad \leq |\{x : \sup_j |(T_j(f-s))(x)| > \varepsilon\}| + 0 + |\{x : \sup_j |(T_j(s-f))(x)| > \varepsilon\}| \\ &\quad = |\{x : (T^*(f-s))(x) > \varepsilon\}| + |\{x : (T^*(s-f))(x) > \varepsilon\}| \\ &\quad \lesssim 2\varepsilon^{-p}\|f - s\|_p^p < 2\varepsilon^{-p}\delta\,.\end{aligned}$$

Since this estimate holds no matter how small δ, we conclude

$$|\{x : |\limsup_{j\to\infty}(T_j f)(x) - \liminf_{j\to\infty}(T_j f)(x)| > 3\varepsilon\}| = 0\,.$$

This concludes the proof of Lemma 16.2 since it shows that the desired limit exists almost everywhere (see also Grafakos [96, Theorem 1.1.11] for the fact that convergence in measure (what we just showed) implies convergence almost everywhere up to a subsequence). □

In the context of these notes, we shall be concerned with the L^p convergence of Bochner–Riesz means. (For pointwise almost everywhere convergence, see, e.g., Carbery [35] and Carbery et al [39] where a maximal weak type (p,p) inequality is proven for $p > 2$, see Tao [184] and Li and Wu [130] for $1 < p < 2$; it is easy to see that $S_R^\delta f$ converges to f uniformly if f is a test function.) By scaling invariance, it suffices to prove the uniform L^p boundedness for $R = 1$. In $d = 1$, it follows from the weak L^1-boundedness of the Hilbert (Riesz) transform and interpolation with the obvious L^2 estimate that S_R is L^p-bounded for all $p \in (1,\infty)$. For $d \geq 2$, one has an explicit kernel representation, namely

Lemma 16.6. *Let $\delta \geq 0$ and $f \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\begin{aligned}(S_1^\delta f)(x) &:= \int_{\mathbb{R}^d} \left(1-|\xi|^2\right)_+^\delta \mathrm{e}^{2\pi i x\cdot\xi} \hat{f}(\xi)\, d\xi = \frac{\Gamma(1+\delta)}{\pi^\delta} \int_{\mathbb{R}^d} \frac{J_{d/2+\delta}(2\pi|x-y|)}{|x-y|^{-d/2-\delta}} f(y)\, dy \\ &\sim \int_{\mathbb{R}^d} \frac{\sum_\pm \mathrm{e}^{\pm 2\pi i|x-y|} + o(1)}{1+|x-y|^{(d+1)/2+\delta}} f(y)\, dy \quad \text{as } |x| \to \infty\,.\end{aligned} \tag{16.2}$$

For a generalization of the $|x| \to \infty$ asymptotics for general $q(\xi)$ (homogeneous of degree one, C^∞, and non-negative in $\mathbb{R}^d \setminus \{0\}$) instead of ξ^2, see [161, Lemma 2.3.3]. Note also that this formula is very similar to the one for $(d\sigma)^\vee$; morally speaking the kernel of $(d\sigma)^\vee$ is comparable to the one of S_1^{-1}. This is akin to the heuristic that the delta function is "of the same strength as" the distribution $1/x$. Note that every time as δ is lowered by 1, (16.2) predicts that the kernel S_1^δ is multiplied by roughly $|x|$. This is consistent with the heuristic observation that the derivative of the symbol $m_\delta = (1-|\xi|^2)_+^\delta$ is roughly comparable to $m_{\delta-1}$.

Proof. See Stein and Weiss [169, Chapter IV, Theorem 4.15], and Tao's notes [179, Lecture 3].

Since the symbol $(1-|\xi|^2)_+^\delta$ is radially symmetric, we only need to compute (with $r = |x|$) the right side of

$$\int_{\mathbb{R}^d} (1-\xi^2)_+^\delta \mathrm{e}^{2\pi i x\cdot\xi}\, d\xi = 2\pi \int_0^\infty (1-k^2)_+^\delta (kr)^{-(d-2)/2} J_{(d-2)/2}(2\pi kr) k^{d-1}\, dk$$

by the Fourier–Bessel transform, see, e.g., Stein and Weiss [169, Chapter IV, Theorem 3.3]. Using the identity

$$J_{\mu+\nu+1}(r) = \frac{r^{\nu+1}}{2^\nu \Gamma(\nu+1)} \int_0^1 J_\mu(kr) k^{\mu+1} (1-k^2)^\nu\, dk$$

for $\mu > -1/2$, $\nu > -1$, and $t > 0$ (see, e.g., Stein–Weiss [169, Chapter IV, Lemma 4.13]), we have (with $\nu = \delta$ and $\mu = (d-2)/2$),

$$\begin{aligned}\int_{\mathbb{R}^d} (1-\xi^2)_+^\delta \mathrm{e}^{2\pi i x\cdot\xi}\, d\xi &= (2\pi)^{1-\delta-1} \cdot 2^\delta \Gamma(1+\delta) r^{-(d-2)/2-\delta-1} J_{(d-2)/2+\delta+1}(2\pi r) \\ &= \frac{\Gamma(1+\delta)}{\pi^\delta} r^{-d/2-\delta} J_{(d-2)/2+\delta+1}(2\pi r)\end{aligned}$$

which yields the first assertion. The asymptotic behavior as $|x| \to \infty$ follows from

$$J_{(d-2)/2+\delta+1}(2\pi r) = \pi^{-1} r^{-1/2} \cos\left(2\pi r - \frac{(d+1+2\delta)\pi}{4}\right) + \mathcal{O}(r^{-3/2})\,, \quad r \to \infty\,,$$

see, e.g., [169, Chapter IV, Lemma 3.11] or Olver [142, Formula 9.2.1]. Note also that the kernel is finite as $|x| \to 0$ since $|J_\nu(x)| \lesssim |x|^\nu$ for $\nu \geq -1/2$, see, e.g., [142, Formula 9.1.62]. (In fact the kernel is complex analytic since the symbol is compactly supported.)

Although the above proof yields the exact formula for the integral kernel, its method is not very robust. Let us therefore now sketch an alternative, somewhat fuzzier, but more robust proof. Since m_δ is radial, we let $\xi = \lambda e_d$ without loss of generality and evaluate in the following

$$\int_{|\xi|\leq 1} (1-\xi^2)^\delta \mathrm{e}^{2\pi i \lambda \xi_d}\, d\xi\,.$$

We decompose this smoothly into three pieces, i.e., the north pole $|\xi - e_n| \ll 1$, the south pole $|\xi + e_n| \gg 1$, and the rest where $|\xi_n| \leq 1 - \varepsilon$ for some $\varepsilon > 0$.

Let's deal with the rest first. By stationary phase, the core part $|\xi| \ll 1$ is rapidly decaying in λ and so it suffices to consider the surface part $|\xi| \sim 1$. In this case, we can use polar coordinates and reduce to

$$\int_{r\sim 1} (1-r^2)_+^\delta r^{d-1} \int_{\mathbb{S}^{d-1}:|\omega_d|\leq 1-\varepsilon} \mathrm{e}^{2\pi i\lambda r\omega_d}\, d\omega\,.$$

But the inner integral is $\mathcal{O}_r(\lambda^{-N})$ for any $N \in \mathbb{N}$ by stationary phase, and so is the total integral.

Thus, we are left to study the north pole (as the south pole is treated analogously). Let us decompose further

$$(1-\xi^2)_+^\delta = f d\omega * d\mu + \text{error}\,,$$

where $f \in C_c^\infty(\mathbb{R}^d)$ is supported on a cap of the north pole, and

$$d\mu(\xi',\xi_d) = \delta(\xi')\eta(\xi_d)(-\xi_d)_+^\delta$$

is a measure supported on the ξ_d axis. Here, $\eta \in C_c^\infty(\mathbb{R})$ is a bump function which equals 1 at the origin. Indeed, one can easily work out that

$$\begin{aligned}(f d\omega * d\mu)(\xi',\xi_d) &= \int_{\mathbb{R}^d} f(\psi',\psi_d)\delta(1-\psi^2)\delta(\xi'-\psi')\eta(\xi_d-\psi_d)(-(\psi_d-\xi_d))_+^\delta\, d\psi \\ &= \int_{\mathbb{R}} f(\xi',\psi_d)\delta(1-\xi'^2-\psi_d^2)\eta(\xi_d-\psi_d)(-(\psi_d-\xi_d))_+^\delta\, d\psi_d \\ &= f(\xi',\Phi(\xi'))\eta(\xi_d-\Phi(\xi'))J(\xi')(\Phi(\xi')-\xi_d)_+^\delta\,,\end{aligned}$$

where J is some Jacobian factor. By choosing f properly, one can make this a good approximation to the kernel of m_δ near the north pole. The error vanishes to order $\delta+1$ or more at the sphere. One can then do a similar decomposition of this error, with a new error term which vanishes to order $\delta+2$. Continuing this procedure shows that one can make the error term as smooth as we like and absorb it into the error term of (16.2).

Let us now consider the contribution of main term, i.e.,

$$(\widehat{f d\omega}\cdot\widehat{d\mu})(\lambda e_d)\,.$$

By the computation of the Fourier transform of surface measures of curved surfaces, the first factor is $C\mathrm{e}^{2\pi i\lambda}\lambda^{-(d-1)/2} + o(\lambda^{-(d-1)/2})$. We claim that the second factor is $(C+o(1))\lambda^{-1-\delta}$. Since the ξ' variable is pretty much irrelevant here, this claim is equivalent to

$$\widehat{d\mu}(\lambda e_n) \sim \mathcal{F}[\eta(\xi_d)(-\xi_d)_+^\delta](\lambda) = (C+o(1))\lambda^{-1-\delta}\,.$$

Recalling that $\mathcal{F}[(\xi_d)_+^\delta](\lambda) = C\lambda^{-1-\delta}$ (in the distributional sense, see, e.g., Gelfand–Shilov [92, Chapter II, Section 2.4 or p. 360]) by homogeneity, the claim follows, since the convolution with the Schwartz function $\hat{\eta}$ does not perturb the decay and merely smoothens out $\mathcal{F}[(\xi_d)_+^\delta](\lambda)$. □

The above representation leads to a necessary condition for the L^p-boundedness of S_R^δ.

Theorem 16.7 (Herz [106])**.** *In order for $\|S_1^\delta f\|_p \lesssim \|f\|_p$ to hold, one must have*

$$|\frac{1}{p}-\frac{1}{2}| < \frac{2\delta+1}{2d}\,. \tag{16.3}$$

In particular, we see that the larger δ gets, the larger the interval on which one has a shot at convergence.

Proof. This is shown by convolving the Bochner–Riesz kernel with a test function of the form

$$f(x) := \begin{cases} 1 & \text{if } |x| < 1/10 \\ 0 & \text{if } |x| \geq 1/10 \end{cases}\,.$$

In this case, $(S_1^\delta f)(x) \sim |x|^{-(d+1+2\delta)/2}$ as $|x| \to \infty$ by Lemma 16.6. Moreover, a moment's thought will convince the reader that the oscillating factor in the Bochner–Riesz kernel produces no significant cancellation in $S_1^\delta f$. Thus, $S_1^\delta f$ does not belong to any L^p if

$$\frac{d}{p} < \frac{d+1}{2} + \delta$$

which is a rearrangement of (16.3). □

We see that the Bochner–Riesz kernel is in L^p for $\delta = 0$ only if $p > 2d/(d+1)$. By duality, it is therefore natural to conjecture that $S_R f$ converges if $p \in (2d/(d+1), 2d/(d-1))$. Let us see whether S_1 is bounded in $d = 1$. In this case

$$S_1 f = \mathcal{F}^{-1}(\mathbf{1}_{[-1,1]}\hat{f}) = \frac{1}{2}\mathcal{F}^{-1}\left((\operatorname{sgn}(x+1) - \operatorname{sgn}(x-1))\hat{f}\right) .$$

By the invariance of multipliers under affine transformations, it thus suffices to prove the L^p-boundedness of $f \mapsto \mathcal{F}^{-1}(\operatorname{sgn}(x)\hat{f})$. But this operator is just the Hilbert transform multiplied by i/π, i.e,

$$\frac{i}{\pi}(Hf)(x) = \frac{i}{\pi}\text{p.v.}\int f(x-y)\frac{dy}{y}$$

which well-known to be L^p-bounded.

Now what about $d \geq 2$? Surprisingly, Fefferman [73] disproved this conjecture, i.e., the ball multiplier is in fact L^p-bounded only for $p = 2$! What is the reason for this dramatic failure of L^p-boundedness?

The previous discussion indicates that Fourier analysis, orthogonality, cancellations and so on should be involved in the analysis of S_R. Fefferman's proof, which was of course via contradiction, involved pure size estimates (he refers to them as Meyer's lemma) and a clever geometric construction. And here is where the Kakeya conjecture comes into the play.

Before we review Fefferman's disproof, let us discuss whether certain regularizations of S_R have a chance of convergence. And indeed, it is conjectured (and in certain cases, such as in $d = 2$, already proven) that the so-called *Bochner–Riesz* means

$$(S_R^\delta f)(x) := \int_{\mathbb{R}^d} \left(1 - |\xi|^2/R^2\right)_+^\delta e^{2\pi i x\cdot\xi}\hat{f}(\xi)\, d\xi$$

do converge for all $\delta > 0$ if p lies in the conjectured range. This the content of

Conjecture 16.8 (Bochner–Riesz conjecture)**.** *Let $\delta > 0$ and $1 \leq p \leq \infty$ be such that* (16.3) *holds. Then $S_R^\delta f$ converges to f in L^p norm as $R \to \infty$ for all $f \in L^p$.*

If p lies outside of the above range, one may still get convergence if δ is chosen to be p-dependent in the right way. As is the case for the restriction conjecture, the Bochner–Riesz conjecture is fully resolved in $d = 2$. (The reason will become clear in a moment.) Observe that for $\delta \to 0$, one recovers the ball multiplier. For higher δ, S_R^δ can indeed be seen as a mollification of the ball multiplier.

Before we continue, let us dispose some easy cases first. Clearly, the conjecture is true when $p = 2$ because of Plancherel's theorem. On the other hand, if $\delta > (d-1)/2$, then, the asymptotics (16.2) of the Bochner–Riesz kernel imply that the convolution kernel S_R^δ is integrable. (Note that there is no singularity near zero in x-space; in fact, $\mathcal{F}[(1-\xi^2)_+^\delta](x)$ must be complex analytic since the multiplier is compactly supported.) So the claim follows by Young's inequality in this case.

Now, what is the connection between the restriction and the Bochner–Riesz conjecture? On the one hand, the implication Restriction ⇒ Bochner–Riesz was shown for the paraboloid by

Carbery [36]. For general surfaces, Fefferman [72] proved that if the (p,p) restriction hypothesis is strengthened to a $(p,2)$ estimate, then the Bochner–Riesz conjecture holds.

On the other hand, the reverse direction Bochner–Riesz $\Rightarrow$ Restriction was shown by Tao [181] for the sphere.

In the following subsections we shall fill in the details in the above discussion. We start by showing the L^p-boundedness of S_1^δ using solely the knowledge of the Bochner–Riesz kernel, Lemma 16.6. Afterwards, we review Fefferman's disproof of the L^p-boundedness of the ball multiplier. We will then review the equivalence Restriction $\Leftrightarrow$ Bochner–Riesz. Finally, we shall see the implication Bochner–Riesz $\Rightarrow$ Kakeya. We will mainly follow [179, Lecture 3], but see also Fefferman [74].

16.1. **L^p-boundedness of S_1^δ via Carleson–Sjölin oscillatory integral estimates.** See also Sogge [161, Section 2.3] for a generalization of the to general $q(\xi)$ (homogeneous of degree one, C^∞, and non-negative in $\mathbb{R}^d \setminus \{0\}$) instead of ξ^2. See also Bourgain [15, 16, 26] and his review [27]. See also Fefferman [74].

From Lemma 16.6 (note that the integral kernel of S_1^δ is complex analytic since the symbol is compactly supported) and Young's inequality it follows immediately that S_1^δ is L^p bounded for all $\delta > (d-1)/2$. The problem gets significantly more difficult in the case $\delta \leq (d-1)/2$ since the kernel is not integrable any more and we need to exploit its oscillatory behavior. Let us recall the necessary condition (Theorem 16.7)

$$\frac{2d}{d-1-2\delta} < p < \frac{2d}{d+1+2\delta}$$

and the known positive result. When $d=2$, matters are completely settled: We shall see below that when $\delta > 0$, S_1^δ is $L^p(\mathbb{R}^2)$ bounded for $4/3 \leq p \leq 4$. There is also the companion result that it actually holds in the range $4/(3+2\delta) < p < 4/(1-2\delta)$, whenever $0 < \delta \leq 1/2$. Our goal in this section is to prove

Proposition 16.9. *The operator S_1^δ, initially defined for $f \in \mathcal{S}$, extends to a $L^p(\mathbb{R}^d)$ bounded operator whenever*

$$\frac{2d}{d+1+2\delta} < p < \frac{2d}{d-1-2\delta}$$

and

$$1 \leq p \leq \frac{2(d+1)}{d+3} \quad \text{or} \quad \frac{2(d+1)}{d-1} \leq p \leq \infty\,.$$

Note that the first restriction is equivalent to

$$\delta > \delta(p) \quad \text{where} \quad \delta(p) = n\left|\frac{1}{p} - \frac{1}{2}\right| - \frac{1}{2}$$

which is the only known necessary condition for the boundedness of S_1^δ. As mentioned above, when $d=2$ this condition is in fact sufficient, i.e., we may drop the second assumption on p. We will prove this fact shortly afterwards.

Let us start with an $L^p \to L^p$ estimate for a certain oscillatory integrals (compare with Theorem 4.3). Let $\psi \in C_c^\infty(\mathbb{R}^d)$ be a smooth cutoff function such that ψ vanishes in a neighborhood of the origin, and set

$$(G_\lambda f)(x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda|x-y|}\psi(x-y)f(y)\,dy\,. \tag{16.4}$$

Invoking the $L^p \to L^q$ Carleson–Sjölin estimates for oscillatory integrals related to the restriction conjecture (see Theorem 4.6 or Theorem A.7) and freezing one variable, we obtain

Lemma 16.10. *We have that*

$$\|G_\lambda f\|_{L^p(\mathbb{R}^d)} \lesssim \lambda^{-d/p'} \|f\|_{L^p(\mathbb{R}^d)} \tag{16.5}$$

whenever $1 \le p \le 2(d+1)/(d+3)$.

Proof. Let us first modify G_λ by setting

$$(\tilde{G}_\lambda f)(x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda|x-y|} \tilde{\psi}(x,y) f(y)\, dy$$

where now $\tilde{\psi} \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ is a smooth cutoff function for $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ whose support does not intersect the diagonal $\{(x,y) : x = y\}$.

For $x = (x', x_d)$, we keep x_d fixed and write

$$(\tilde{G}_\lambda f)(x', x_d) = (T_\lambda^* f)(x')$$

where

$$(T_\lambda^* f)(x') = \int_{\mathbb{R}^d} \mathrm{e}^{-i\lambda\varphi(x',y)} \overline{\tilde{\psi}}(x', y) f(y)\, dy\,.$$

(i.e., the restriction operator, recall also Theorem 4.6 and Appendix A.2). This leads us to the phase function $\varphi(x', y)$ on $\mathbb{R}^{d-1} \times \mathbb{R}^d$ given by

$$\varphi(x', y) = -(|x' - y'|^2 + |x_d - y_d|^2)^{1/2}\,,$$

with x_d fixed and $y = (y', y_d)$. It is not difficult to verify directly that φ satisfies the conditions of Theorem A.7. Indeed, the vector $\overline{u}$ arising in the curvature hypothesis (A.11) may be taken to be $\overline{u} = (x-y)/|x-y|$. We can therefore invoke Theorem A.7 and obtain

$$\left(\int_{\mathbb{R}^{d-1}} |(\tilde{G}_\lambda f)(x', x_n)|^q \, dx' \right)^{1/q} \lesssim \lambda^{-d/p'} \|f\|_{L^p(\mathbb{R}^d)}\,.$$

Next observe that $q \ge p$ and that the integration in x' above is only over a compact set. Thus,

$$\int_{\mathbb{R}^{d-1}} |(\tilde{G}_\lambda f)(x', x_d)|^p \, dx' \lesssim \lambda^{-dp/p'} \|f\|^p_{L^p(\mathbb{R}^d)}$$

and a final integration in x_d (again over a compact set) gives

$$\|\tilde{G}_\lambda f\|_{L^p(\mathbb{R}^d)} \lesssim \lambda^{-d/p'} \|f\|_{L^p(\mathbb{R}^d)}\,.$$

The passage from the inequality for $\tilde{G}_\lambda$ to that for G_λ (i.e., to go back to a $C_c^\infty(\mathbb{R}^d)$ function ψ from a $C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ function $\tilde{\psi}$) is then accomplished by a familiar argument, see, e.g., Stein [167, Chapter VI, Section §2.3]. Indeed, the last estimate implies

$$\int\limits_{|x-x^0|\le 1} |(G_\lambda f)(x)|^p \, dx \lesssim \lambda^{-dp/p'} \int\limits_{|x-x^0|\le c} |f(x)|^p \, dx$$

for each x^0, where the constant c is determined by the size of the support of ψ. An integration in $x^0 \in \mathbb{R}^d$ (which only yields a multiple of the volume of the unit ball) then proves the assertion of the lemma. ☐

Proof of Proposition 16.9. Recall that $S_1^\delta f = K_\delta * f$ with K_δ as in Lemma 16.6 where the principal term is given by a constant multiple of

$$\int\limits_{|y|\ge 1} \mathrm{e}^{\pm 2\pi i|y|} f(x-y)|y|^{-(d+1)/2-\delta} \, dy \equiv (Tf)(x)\,.$$

Then there are finitely many terms of the same kind, but where the factor $|y|^{-(d+1)/2-\delta}$ is replaced by $|y|^{-(d+1)/2-\delta-j}$ (and hence improved) with $j > 0$. Finally there is an error term which corresponds to the convolution with an L^1 kernel. Thus, we only need to deal with the principal term.

Let us now decompose

$$|y|^{-(d+1)/2-\delta} = \sum_{k\geq 0} 2^{-[(d+1)/2+\delta]k} \cdot \left(\frac{|y|}{2^k}\right)^{-(d+1)/2-\delta} \psi\left(\frac{y}{2^k}\right)$$

dyadically where (as before) $\psi(x) = \varphi(x) - \varphi(2x)$ is a smooth function supported in $1/2 < |x| < 2$ (when φ is a bump function at the origin). Thus, we may write $T = \sum_{k\geq 0} T_k$ where

$$(T_k f)(x) = 2^{-[(d+1)/2+\delta]k} \int_{\mathbb{R}^d} \mathrm{e}^{2\pi i|y|} f(x-y) \left(\frac{|y|}{2^k}\right)^{-(d+1)/2-\delta} \psi\left(\frac{y}{2^k}\right) .$$

Now, scaling $y \mapsto 2^k y$ shows that, whenever $1 \leq p \leq 2(d+1)/(d+3)$,

$$\|T_k\|_{p,p} = 2^{-[(d+1)/2+\delta]k} \|G_{2\pi\cdot 2^k}\|_{p,p} \cdot 2^{dk} \lesssim 2^{-[(d+1)/2+\delta]k} \cdot 2^{-dk/p'} \cdot 2^{dk}$$

with G as in the previous lemma where $\psi(y)$ is replaced by $|y|^{-(d+1)/2-\delta}\psi(y)$. If

$$-\left[\frac{d+1}{2} + \delta\right] - \frac{d}{p'} + d < 0 \,,$$

which is equivalent to $p > 2d/(d+1+2\delta)$ (i.e., the asserted range for p), then $\|T\|_{p,p} \lesssim \sum_{k\geq 0} \|T_k\|_{p,p}$ converges which concludes the proof. □

We shall now review Carleson's and Sjölin's proof [44] of the Bochner–Riesz conjecture in $d = 2$. We emphasize that the following estimate extends Theorem A.7 to the full range $1 \leq p < 4$ (instead of $1 \leq p \leq 2$).

Theorem 16.11. *Under the assumptions of Theorem A.7, when $d = 2$, we have*

$$\|T_\lambda f\|_{L^q(\mathbb{R}^2)} \lesssim \lambda^{-2/q} \|f\|_{L^p(\mathbb{R}^1)} \tag{16.6}$$

where $q = 3p'$ and $1 \leq p < 4$.

Proof. See [167, p. 412]. □

As a corollary, one obtains the full Bochner–Riesz and restriction conjectures in $d = 2$. The latter is essentially contained in Fefferman [72] (joint with E. M. Stein). For an alternative proof of the Bochner–Riesz conjecture in $d = 2$, see also Fefferman [74].

Corollary 16.12. *Suppose $S \subseteq \mathbb{R}^2$ is a curve whose curvature is nowhere zero and S_0 is a compact subset of S. Then*

$$\left(\int_{S_0} |\hat{f}(\xi)|^q \, d\sigma(\xi)\right)^{1/q} \lesssim_{S_0} \|f\|_{L^p(\mathbb{R}^2)} \,, f \in \mathcal{S} \,,$$

whenever $3q = p'$ and $1 \leq p < 4/3$.

For an alternative proof of this, we refer to Subsection 7.4 (which followed [179, Lecture 5]).

Corollary 16.13. *The operator S_1^δ extends to a $L^p(\mathbb{R}^2)$ bounded operator for $4/3 \leq p \leq 4$ whenever $\delta > 0$ and more generally to the range*

$$\frac{4}{3+2\delta} < p < \frac{4}{1-2\delta} \,,$$

whenever $0 < \delta \leq 1/2$.

16.2. **The multiplier problem for the ball.** We review Fefferman's disproof of the boundedness of the disk multiplier using a variant of the Kakeya conjecture [73]. Nice expositions can also be found in Krantz [126, Section 3.5] and Grafakos [97, Section 5.1].

As we have already mentioned several times, Carleson and Sjölin [44] made heartening progress in 1972 when they proved that the disc multiplier is almost L^p bounded in the sense that S_1^δ is L^p bounded for any $\delta > 0$ and $4/3 \leq p \leq 4$ using the theory of oscillatory integrals. In this section, we shall show that this is indeed the best that one can get. Writing $S^\delta \equiv S_1^\delta$, we show

Theorem 16.14. *S^0 is bounded only in $L^2(\mathbb{R}^d)$ for $d \geq 2$.*

Indeed it suffices to disprove L^p boundedness for $p > 2$ (by duality, we also obtain the case $p < 2$) in two dimensions since L^p boundedness in $\mathbb{R}^d$ implies boundedness in $\mathbb{R}^{d-1}$ by an observation of de Leeuw.

Lemma 16.15 (de Leeuw)**.** *Suppose that m is a smooth Fourier multiplier on $\mathbb{R}^d$ and that the operator T defined by*

$$\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$$

is bounded on $L^p(\mathbb{R}^d)$. Then the operator T_0 defined by

$$\widehat{T_0 g}(\xi') = m(\xi', 0)\hat{g}(\xi')$$

for $\xi' \in \mathbb{R}^{d-1}$ is bounded on $L^p(\mathbb{R}^{d-1})$.

Proof. From the invariance of L^p multiplier bounds under affine transformations (to see this, just scale), we see that we may replace $m(\xi)$ by

$$m_R(\xi', \xi_d) = m(\xi', \xi_d/R)$$

in the definition of T without affecting the L^p boundedness property. Letting $R \to \infty$ and taking limits, we may replace m by

$$m_\infty(\xi', \xi_d) = m(\xi', 0)\,,$$

i.e., the operator

$$\widehat{T_\infty f}(\xi', \xi_d) = m(\xi', 0)\hat{f}(\xi', \xi_d)$$

is bounded on $L^p(\mathbb{R}^d)$. If we now apply this fact to a function of the form $f(x', x_d) = g(x')\psi(x_d)$ and observe that $\hat{f}(\xi', \xi_d) = \hat{g}(\xi')\hat{\psi}(\xi_d)$, we obtain the desired result. □

There are two key insights in the disproof of the disc conjecture. The first is that the disc conjecture would imply a vastly improved Kakeya conjecture (Meyer's lemma) where the "tubes" will not have to be separated anymore. The second key is that such a strengthened Kakeya estimate can indeed never hold. The proof of the latter is inspired by Besicovitch's (or rather Schönberg's simplified) construction of Besicovitch sets (that contain a unit line segment in every direction). Let us start with the first insight.

Lemma 16.16 (Y. Meyer)**.** *Let $(v_j)_{j\in\mathbb{N}}$ be a sequence of unit vectors in $\mathbb{R}^2$ and let H_j be the half-plane $\{x \in \mathbb{R}^2 : x \cdot v_j \geq 0\}$. Defined the "half plane multipliers" $(T_j)_{j\in\mathbb{N}}$ on $L^p(\mathbb{R}^2)$ by setting $\widehat{T_j f}(\xi) = \mathbf{1}_{H_j}(\xi)\hat{f}(\xi)$. If the disc conjecture holds, then for any sequence $(f_j)_{j\in\mathbb{N}}$, we have the square function estimate*

$$\|(\sum_j |T_j f_j|^2)^{1/2}\|_p \lesssim \|(\sum_j |f_j|^2)^{1/2}\|_p\,. \tag{16.7}$$

Proof. The idea is to approximate the half-planes by gigantic discs and to use the standard randomization argument to obtain the above square function estimate from the supposed L^p boundedness of the disc multiplier. More precisely, let $T_{D_j^r}$ be the operator defined by $\widehat{T_{D_j^r} f}(\xi) =$

$\mathbf{1}_{D_j^r}\hat{f}(\xi)$ where D_j^r is the disc of radius r centered at rv_j. For $f \in C_c^\infty$, we have the uniform convergence

$$(T_j f)(x) = \lim_{r\to\infty} (T_{D_j^r} f)(x)$$

which is easy by going to Fourier space since

$$\|(\mathbf{1}_{H_j} - \mathbf{1}_{D_j^r})\hat{f}\|_1 \le \|(\mathbf{1}_{H_j} - \mathbf{1}_{D_j^r})\|_\infty \|\hat{f}\|_1 \to 0\,.$$

Thus, by Fatou's lemma

$$\|(\sum_j |T_j f_j|^2)^{1/2}\|_p \le \liminf_{r\to\infty} \|(\sum_j |T_{D_j^r} f_j|^2)^{1/2}\|_p$$

By dilating $\mathbb{R}^2$ it therefore suffices to set $r = 1$ and prove

$$\|(\sum_j |T_{D_j^1} f_j|^2)^{1/2}\|_p \lesssim \|(\sum_j |f_j|^2)^{1/2}\|_p\,.$$

Since translating in Fourier space corresponds to multiplying by phases in position space, we have (recalling that S^0 was the disc multiplier)

$$(T_{D_j^1} f)(x) = \mathrm{e}^{2\pi i v_j\cdot x} S^0[\mathrm{e}^{-2\pi i v_j\cdot y} f](x)\,,$$

and so it suffices to prove

$$\|(\sum_j |S^0[\mathrm{e}^{-2\pi i v_j\cdot y} f_j]|^2)^{1/2}\|_p \lesssim \|(\sum_j |f_j|^2)^{1/2}\|_p\,.$$

But by the Marcinkiewicz–Zygmund theorem (see, e.g., Grafakos [96, Theorem 5.5.1]), this estimate holds because of the assumed L^p boundedness of S^0. □

To disprove the disc conjecture, we shall find a counterexample to the square function estimate (16.7) for half planes. The example is based on a slight variant of (Schönberg's improvement of) Besicovitch's construction for the Kakeya needle problem.

Lemma 16.17. *Fix a small number $\eta > 0$. Then there is a set $E \subseteq \mathbb{R}^2$ and a collection $\mathcal{R} = \{R_j\}_{j\in\mathbb{N}}$ of pairwise disjoint rectangles with the properties that*

(1) $|E \cap \tilde{R}_j| \ge |\tilde{R}_j|/10$, i.e., at least one-tenth of the area of each $\tilde{R}_j$ lies in E and
(2) $|E| \le \eta \sum_j |R_j|$

where $\tilde{R}_j$ is the shaded region in Figure 6.

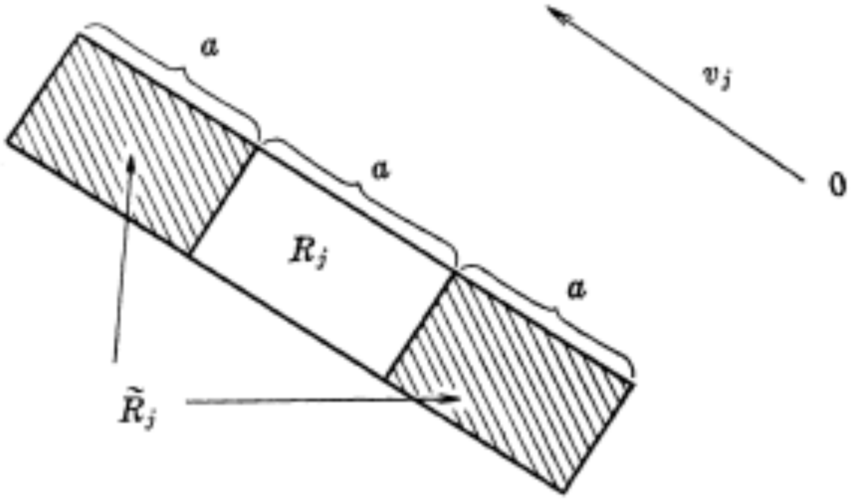

Figure 6

Let us now see how the half-plane multiplier acts on functions supported on rectangles whose long side is oriented along the normal of the half plane.

Lemma 16.18. *Let R be a $a \times b$ rectangle in the plane with arbitrary position and orientation and let $\tilde{R}$ be the rectangle of the same length which is shifted over by $c \cdot a$ for some constant $c \geq 1$ in the direction of the long axis of R. Then there exists a function f_R supported on R such that $|f_R| \leq 1$ on R and $|(T_j f_R)(\tilde{x})| \sim 1$ for any $v_j \in \mathbb{S}^1$ and $\tilde{x} \in \tilde{R}$.*

Observe that for $c = 1$, we recover the setup of Figure 6. The following arguments can easily be generalized to treat also the case $0 < c < 1$, which is left as an exercise.

Proof. Let us assume $v_j = (-1, 0)$, i.e., we consider the half-plane $\xi_1 \leq 0$, i.e., $\widehat{T_j f}(\xi) \equiv \chi(\xi)\hat{f}(\xi)$ where

$$\chi(\xi) = \frac{1}{2} - \frac{1}{2}\operatorname{sgn}(\xi_1)\,.$$

By the formula for the Fourier transform of the Hilbert transform, we have (in the sense of distributions)

$$(T_j f)(x) = \int_{\mathbb{R}^2} \left(\frac{1}{2}\delta(y) + \frac{1}{2\pi i}\frac{1}{y_1}\delta(y_2)\right) f(y - x)\, dy\,.$$

Now, let $0 \leq \psi \in C_c^\infty(\mathbb{R}^2)$ be supported on $[0,1]^2$ with $\psi > 0$ on $[1/3, 2/3]^2$ and symmetric with respect to reflections with respect to the coordinate axes. Let furthermore R be the rectangle centered at $z \in \mathbb{R}^2$ whose long side a lies in the x_1 direction and whose short side b lies along the x_2 direction. If we define

$$f_R(x) = \psi\left(\frac{x_1 - z_2}{a}, \frac{x_2 - z_2}{b}\right)\,,$$

then the action of the half-plane multiplier on f_R evaluated at the center $\tilde{x} = (z_1 + ca, z_2)$ of the translated $\tilde{R}$ is given by

$$\begin{aligned}(T_j f_R)(\tilde{x}) &= \frac{1}{2}\int_{\mathbb{R}^2}\left(\delta(y) + \frac{1}{i\pi}\delta(y_2)y_1^{-1}\right)\psi\left(y_1 - \frac{\tilde{x}_1 - z_1}{a}, y_2 - \frac{\tilde{x}_2 - z_2}{b}\right)\,dy \\ &= \frac{1}{2}\psi\left(c, \frac{\tilde{x}_2 - z_2}{b}\right) + \frac{1}{2\pi i}\int_{\mathbb{R}}\frac{dy_1}{y_1}\psi\left(y_1 - c, \frac{\tilde{x}_2 - z_2}{b}\right).\end{aligned}$$

The first summand vanishes since $c > 1$, whereas the second one, in absolute modulus at least, is clearly bounded from below by some positive constant. The computation where $\tilde{x}$ is an arbitrary point in $\tilde{R}$ is completely analogous. □

With this at hand, we can disprove the disc conjecture by contradicting the square function estimate (16.7) for half-planes.

Proof of Theorem 16.14. Set $f_j = \mathbf{1}_{R_j}$ with R_j as in Figure 6 and v_j being parallel to the longer sides of R_j. Direct computation shows that $|(T_j f_j)(x)| \geq 1/2$ for $x \in \tilde{R}_j$, so that

$$\int_E (\sum_j |(T_j f_j)(x)|^2)\,dx = \sum_j \int_E |(T_j f_j)(x)|^2\,dx \geq \frac{1}{4}\sum_j |E \cap \tilde{R}_j| \geq \frac{1}{40}\sum_j |\tilde{R}_j| = \frac{1}{20}\sum_j |R_j| \tag{16.8}$$

by the fact $|E \cap \tilde{R}_j| \geq |\tilde{R}_j|/10$ for our constructed set E. On the other hand, if the square function estimate (16.7) were true, Hölder's inequality would show that the left side of (16.8) is bounded from above by

$$\begin{aligned}\int_E (\sum_j |(T_j f_j)(x)|^2)\,dx &\leq |E|^{(p-2)/p}\ \|(\sum_j |T_j f_j|^2)\|_p^2 \lesssim |E|^{(p-2)/p}\ \|(\sum_j |f_j|^2)\|_p^2 \\ &= |E|^{(p-2)/p}\ (\sum_j |R_j|)^{2/p} \leq \eta^{(p-2)/p}\sum_j |R_j|\end{aligned} \tag{16.9}$$

where we first used the square function estimate, then the fact that the R_j are pairwise disjoint, i.e., there are no mixed terms appearing in the summation over j, and finally the size assumption $|E| \leq \eta\sum_j |R_j|$ on the constructed set E. For sufficiently small η the bounds in (16.8) and (16.9) contradict each other which disproves the square function estimate (16.7). This shows the failure of the L^p boundedness of the disc multiplier and concludes the proof of Theorem 16.14. □

We are thus left to give the

Proof of Lemma 16.17. We shall closely follow the excellent exposition of Cunningham [59] (where the minimal area for a plane, simply connected, or star-shaped, set within which a unit segment can be rotated continuously to return to its original position with its ends reversed, is determined; in fact, it is shown that star-shaped Kakeya sets *cannot* have area less than $\pi/108$, although it was not known whether this is the best value), but see also the classic paper of Busemann and Feller [34].

Consider the following process: we are given a triangle T as in the left drawing in Figure 7, with horizontal base ab and height h. Extend the lines ac and bc to points a' and b' of height $h' > h$. Let d be the midpoint of ab, see the right drawing in Figure 7.

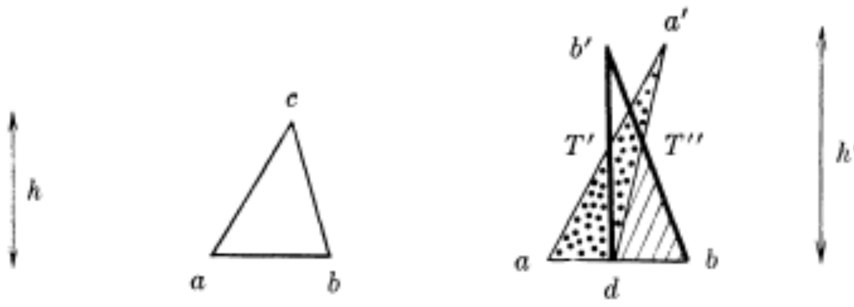

Figure 7

We say that the two triangles $T' = ada'$ and $T'' = bdb'$ arise as *sprouts* from height h to height h'.

Now we can construct the Besicovitch set E. Begin with an equilateral triangle T^0 whose base is the interval $[0, 1]$ on the x-axis, and pick an increasing sequence of numbers $h_0, h_1, h_2, ..., h_k$, where $h_0 = \sqrt{3}/2$ denotes the height of the initial triangle T^0. Now sprout T^0 from height h_0 to height h_1 to obtain two new triangles T' and T''. Now sprout both T' and T'' from height h_1 to height h_2 to obtain four new triangles T^1, T^2, T^3, T^4, all of height h_2. Continue sprouting, obtaining at stage n, 2^n triangles of height h_n with base length 2^{-n}. Finally, set E equal to the union the final 2^k triangles $T^1, T^2, ..., T^{2^k}$ which arose at stage k.

For the special case, where $h_0 = \sqrt{3}/2$, we obtain the sequence of heights

$$h_0 = \frac{\sqrt{3}}{2}\,, h_1 = \frac{\sqrt{3}}{2}\left(1 + \frac{1}{2}\right)\,, h_2 = \frac{\sqrt{3}}{2}\left(1 + \frac{1}{2} + \frac{1}{3}\right)\,, ..., h_k = h_1 = \frac{\sqrt{3}}{2}\sum_{n=1}^{k} n^{-1} \sim \log k$$

Buseman and Feller [34] showed that $|E| \le 17$. (Actually, Busemann and Feller use a sprouting procedure slightly different from this. However, since their sprouted triangles are strictly larger than these, their estimates apply here, too.)

Having built E and computed its measure, we are left to construct the collection of disjoint rectangles which satisfied $|E \cap \tilde{R}_j| \ge |\tilde{R}_j|/10$ and $|E| \le \eta \sum_j |R_j|$ for any given (small) $\eta > 0$. To do so note that each dyadic interval $I \subseteq [0, 1]$, of length 2^{-k}, is the base of *exactly* one $T_j =: T(I)$. Let us call its upper vertex $P(I)$. We then construct the rectangle $R(I)$ as in Figure 8.

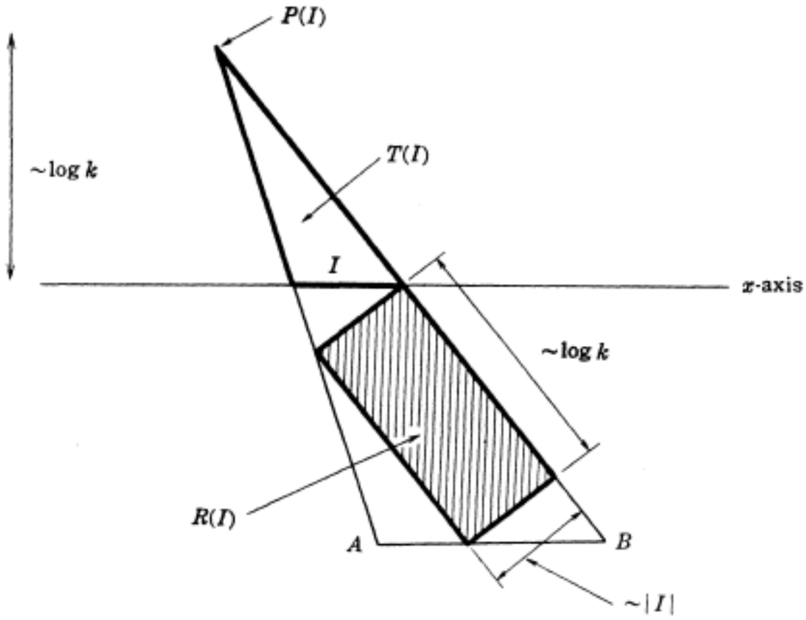

Figure 8

It does not matter how $R(I)$ is placed, as long as it stays inside the triangle $P(I)BA$. Now define

$$\mathcal{R} = \{R(I) : I \text{ is a dyadic subinterval of } [0,1] \text{ of length } 2^{-k}\} .$$

Now let us check the claimed properties. First, $|E \cap \tilde{R}_j| \geq |\tilde{R}_j|/10$ is trivially satisfied by construction since $T(I) \subseteq E$. To check the upper bound $|E| \leq \eta \sum_I |R_I|$, we note that the area of each $R(I)$ is roughly $2^{-k} \log k$. Since there are altogether 2^k of such rectangles, we have

$$\sum_I |R_I| \sim 2^k \cdot 2^{-k} \log k = \log k .$$

Clearly, the left side is greater than $|E|/\eta$ if we pick $k = k(\eta)$ so large that $\log k > 17/\eta$.

Finally, it remains to show that the rectangles are pairwise disjoint. But this just follows from the elementary geometric observation that $P(I')$ lies to the left of $P(I)$ whenever I' lies to the left of I. □

16.3. **Restriction ⇒ Bochner–Riesz.** This is essentially contained in Fefferman [72, Theorem 3] but we will follow the exposition in [179, Lecture 3].

Let us fix $\delta > 0$ such that the necessary condition (16.3) holds. Then, as in the proof of the Tomas–Stein theorem, we will decompose the convolution kernel $K_\delta = \mathcal{F}[(1-\xi^2)_+^\delta]$ dyadically using the $\psi_k(x) := \varphi(2^{-k}x) - \varphi(2^{-k+1}x)$ where φ was a bump function supported around the origin. Then, we break up

$$K_\delta = \varphi K_\delta + \sum_{k>0} \psi_k K_\delta .$$

As opposed to the proof of the Tomas–Stein theorem, we do not need to impose any fancy moment conditions on φ or ψ since we inequality on p is strict, i.e., we do not need to care about any subtleties concerning endpoints.

First, since φK_δ is a bump function, the convolution is clearly an L^p-bounded operator by Young's inequality. So, as before, we are left with showing that

$$\| \sum_k f * (\psi_k K_\delta) \|_p \lesssim \|f\|_p \,.$$

Since we have a bit of room in the condition (16.3) on p, we may just use the triangle inequality. In fact, we shall show

$$\|f * (\psi_k K_\delta)\|_p \lesssim 2^{[d(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}-\delta]k} \|f\|_p \tag{16.10}$$

which is just summable in k if (16.3) holds.

The first observation that we shall use to prove (16.10) is that the kernel $\psi_k K_\delta$ is compactly supported on an annulus $\{x : |x| \sim 2^k\}$, i.e., the operator is *somewhat localized.* In fact, the values of f at a point x only influence points which are in a 2^k neighborhood. The following useful lemma allows us to reduce our study of such "local" operators to a compact set.

Lemma 16.19. *Let T be a linear operator taking functions on $\mathbb{R}^d$ to functions on $\mathbb{R}^d$. Suppose T is local in the sense that the support of Tf always remains within R of the support of f for some $R > 0$. Then, for any $1 \le p \le q \le \infty$, the bound*

$$\|Tf\|_q \lesssim \|f\|_p \quad \text{for all } f \in L^p(\mathbb{R}^d) \tag{16.11}$$

is equivalent to the bound

$$\|Tf\|_{L^q(B_x(2R))} \lesssim \|f\|_p \quad \text{for all } f \in L^p(B_x(R))\,, \tag{16.12}$$

holding uniformly in x.

In other words, to show (16.11), it suffices to test it for functions supported on an R-ball. Intuitively, the idea is that functions on *distinct R-balls* basically do not interfere too much with each other.

Proof. Clearly, we only need to show (16.12) $\Rightarrow$ (16.11). For this purpose let $f \in L^p(\mathbb{R}^d)$, choose a finitely overlapping collection of balls $\{B\}$ that cover $\mathbb{R}^d$ and denote a partition of unity $1 = \sum_B \psi_B$ subordinate to that cover. Then, we write

$$\|Tf\|_q^q = \int |T(\sum_B \psi_B f)|^q = \int |\sum_B T(\psi_B f)|^q \,.$$

Since T is local in the above sense, the functions $T(\psi_B f)$ are just supported on the double $2B$ of B. These balls are still only finitely overlapping, so we have the pointwise estimate

$$|\sum_B T(\psi_B f)|^q \lesssim \sum_B |T(\psi_B f)|^q \,.$$

Putting this back into the previous estimate, simplifying, applying the assumed $L^p \to L^q$-boundedness (16.12), and the elementary inequality

$$\left(\sum_B a_B^q\right)^{1/q} \le \left(\sum_B a_B^p\right)^{1/p}$$

for a sequence $\{a_B\}_B$ of non-negative numbers and (crucially) $q \ge p$, we obtain

$$\|Tf\|_q \lesssim \left(\sum_B \|T(\psi_B f)\|_q^q\right)^{1/q} \lesssim \left(\sum_B \|\psi_B f\|_p^q\right)^{1/q} \le \left(\sum_B \|\psi_B f\|_p^p\right)^{1/p} .$$

Again, since the balls are only finitely overlapping, it is easy to see that the right side is essentially $\|f\|_p$. $\square$

The condition $q \geq p$ in the above lemma is absolutely necessary. This is an example of one of *Littlewood's principles*: "the higher exponents are always to the left". More precisely, we have

Lemma 16.20. *Let $1 \leq p, q < \infty$ and T be a non-zero translation invariant operator on $\mathbb{R}^d$. Then, the estimate $\|Tf\|_q \lesssim \|f\|_p$ is only possible, if $q \geq p$.*

Proof. Let φ be any bump function such that $T\varphi$ is non-zero. Let $N > 0$ be a large number, and let $x_1, ..., x_N$ be N very widely separated points. Define f by

$$f(x) = \sum_{i=1}^{N} \varphi(x - x_i) \,.$$

If the above estimate held for f, we would have

$$\| \sum_{i=1}^{N} T\varphi(x - x_i)\|_q \lesssim \| \sum_{i=1}^{N} \varphi(x - x_i)\|_p$$

since T is translation invariant. However, the right side is bounded from above by a constant ($\|\varphi\|_p$) times $N^{1/p}$, whereas the left side can in fact be bounded from below (by forgetting about the overlaps of $T\varphi(x - x_i)$) by a constant ($\|T\varphi\|_q$) times $N^{1/q}$. Letting $N \to \infty$, we have necessarily $1/q \leq 1/p$, i.e., $q \geq p$. $\square$

Now, let us return to the proof of (16.10), i.e.,

$$\|f * (\psi_k K_\delta)\|_p \lesssim 2^{[d(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}-\delta]k} \|f\|_p \,.$$

From the above discussion, it suffices to prove

$$\|f * (\psi_k K_\delta)\|_{L^p(B_x(a2^k))} \lesssim 2^{[d(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}-\delta]k} \|f\|_p$$

for all f supported on a ball $B_x(a2^{k-1})$. By translation invariance, we may take $x = 0$.

We are supposed to apply the Tomas–Stein theorem which is an $L^p \to L^2$ theorem. Indeed, using Hölder's inequality (since we are on a finite domain) on the left side of the last formula and using Plancherel, we have

$$\|f * (\psi_k K_\delta)\|_{L^p(B_x(a2^k))} \lesssim 2^{dk(\frac{1}{p}-\frac{1}{2})} \|f * (\psi_k K_\delta)\|_2 = 2^{dk(\frac{1}{p}-\frac{1}{2})} \|\hat{f} \cdot (\widehat{\psi_k} * m_\delta)\|_2$$

where we have denoted $m_\delta(\xi) = (1 - \xi^2)_+^\delta$. Thus, we are left to show

$$\|\hat{f} \cdot (\widehat{\psi_k} * m_\delta)\|_2 \lesssim 2^{-(\frac{1}{2}+\delta)k} \|f\|_p \,.$$

We will shortly prove the key estimate

$$|\hat{\psi}_k * m_\delta(\xi)| \lesssim 2^{-\delta k}(1 + 2^k d(\xi, S))^{-N} \,, \quad N \in \mathbb{N} \,. \tag{16.13}$$

Assuming this for a moment, we see that it suffices to prove

$$\|\hat{f}(1 + 2^k d(\xi, S))^{-N}\|_2^2 = \int \frac{|\hat{f}(\xi)|^2}{(1 + 2^k d(\xi, S))^{2N}} \, d\xi \lesssim 2^{-k} \|f\|_p^2$$

to finish the proof. We distinguish between $d(\xi, S) > 1/2$ and $d(\xi, S) < 1/2$ and start with the former case, which is an error term. In this case, we crudely estimate

$$|\hat{f}| \lesssim 2^{ak} \|f\|_p$$

by the definition of $\hat{f}$, Hölder's inequality, and the fact that f is compactly supported. On the other hand, the denominator in the integral is 2^{-Nk} for any N and rapidly decreasing as $\xi \to \infty$.

This decay beats all other factors, and the bound is easy to prove. Thus, it suffices to prove

$$\int\limits_{1/2\leq|\xi|\leq 3/2} \frac{|\hat{f}(\xi)|^2}{(1+2^k d(\xi,S))^{2N}} \lesssim 2^{-k}\|f\|_p^2\,.$$

Discarding the Jacobian arising from passing to polar coordinates, we rewrite this as

$$\int_{1/2}^{3/2} dr\ (1+2^k|r-1|)^{-2N} \int_{r\mathbb{S}^{d-1}} d\omega\ |\hat{f}|^2 \lesssim 2^{-k}\|f\|_p^2\,.$$

16.4. **Bochner–Riesz ⇒ Restriction.** We review Tao's proof [181] that the Bochner–Riesz conjecture implies the restriction conjecture.

16.5. **Bochner–Riesz ⇒ Kakeya.** We review the argument Bochner–Riesz ⇒ Kakeya. We discuss Bourgain's works [15, 26, 17] that progress on Kakeya is connected to progress for Bochner–Riesz (and thereby for Restriction by Tao [181])

16.6. **How does Kakeya help in proving Bochner–Riesz?** The key observation is that every function can be decomposed into a linear combination of wave packets by applying standard cutoffs both in physical space (by pointwise multiplication) and in frequency space (using the Fourier transform). After applying the Bochner–Riesz operator to these wave packets individually, one has to reassemble the wave packets and obtain estimates for the sum. Kakeya estimates play an important role in this since the wave packets are essentially supported on tubes; however, this is not the full story since these packets also carry some oscillation that can be exploited. Thus, one must develop tools to deal with the possible cancellation between wave packets. The known techniques to deal with this cancellation, mostly based on L^2 methods, are imperfect, so that even if one had a complete solution to the Kakeya conjecture, one could not then completely solve the Bochner–Riesz conjecture. Nevertheless, the best-known results on Bochner–Riesz (e.g., in $d = 3$ the conjecture is known, see Tao and Vargas [177, 178] for $p > 26/7$ and for $p < 26/19$ using also bilinear methods) have been obtained by utilizing the best-known quantitative estimates of Kakeya type.

17. Connection to spectral multipliers

17.1. **Eigenfunction estimates for $-\Delta$.** We start with the basic observation

$$dE_{\sqrt{-\Delta}}(\lambda) = \lambda^{d-1} R_{\lambda\mathbb{S}^{d-1}}^* R_{\lambda\mathbb{S}^{d-1}}\, d\lambda$$

in the sense that for $f \in \mathcal{S}(\mathbb{R}^d)$,

$$dE_{\sqrt{-\Delta}}(\lambda)f(x) = \lambda^{d-1}\int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi i x\cdot(\lambda\omega)}\hat{f}(\lambda\omega)\,d\sigma(\omega) = \int_{\lambda\mathbb{S}^{d-1}} \mathrm{e}^{2\pi i x\cdot\xi}\hat{f}(\xi)\,d\sigma_{\lambda\mathbb{S}^{d-1}}(\xi)\,.$$

Here $dE_A(\lambda)$ denotes the spectral projection associated to some self-adjoint operator A. This follows immediately from

$$\int_0^\infty F(\lambda)\langle\psi, dE_{\sqrt{-\Delta}}(\lambda)\psi\rangle = \langle\psi, F(\sqrt{-\Delta})\psi\rangle = \int_0^\infty dk\ F(k)\cdot\left(k^{d-1}\int_{\mathbb{S}^{d-1}}|\hat{\psi}(k\omega)|^2\,d\sigma(\omega)\right)$$

for appropriate measurable functions $F : [0,\infty) \to \mathbb{R}$. In particular, the (rescaled) Tomas–Stein estimate

$$\begin{aligned} k^{d-1}\int_{\mathbb{S}^{d-1}} |\underbrace{\hat{f}(k\omega)}_{=:\hat{g}_k(\omega)}|^2\,d\omega &\lesssim k^{d-1}\|g_k\|_{p_c}^2 = k^{d-1}\|k^{-d}f(\cdot/k)\|_{p_c}^2 = k^{d-1-2d+2d/p_c}\|f\|_{p_c}^2 \\ &= k^{-d+2d/p_c-1}\|f\|_{p_c}^2\,, \end{aligned}$$

with $p_c = 2(d+1)/(d+3)$ immediately yields

$$\left\| \frac{dE_{\sqrt{-\Delta}}(\lambda)}{d\lambda} \right\|_{p_c \to p_c'} \lesssim \lambda^{-d+2d/p_c-1} = \lambda^{d(\frac{1}{p_c}-\frac{1}{p_c'})-1} \tag{17.1}$$

By a change of variables and the rescaled Tomas–Stein estimate, i.e.,

$$\begin{aligned}\int_0^\infty F(\lambda)\langle \psi, dE_{-\Delta}(\lambda)\psi\rangle = \langle \psi, F(-\Delta)\psi\rangle &= \int_0^\infty dk\ F(k^2)\cdot \left(k^{d-1}\int_{\mathbb{S}^{d-1}} |\hat\psi(k\omega)|^2\, d\sigma(\omega)\right)\\ &= \frac{1}{2}\int_0^\infty dk\ F(k)\cdot\left(k^{d/2-1/2-1/2}\int_{\mathbb{S}^{d-1}}|\hat\psi(\sqrt{k}\omega)|^2\, d\sigma(\omega)\right)\\ &\lesssim \int_0^\infty dk\ F(k)\cdot k^{d/2-1-d+d/p_c}\|\psi\|_{L^{p_c}}^2\end{aligned} \tag{17.2}$$

for any $F:[0,\infty)\to[0,\infty)$ (such as a characteristic function), we obtain analogously

$$\left\| \frac{dE_{-\Delta}(\lambda)}{d\lambda} \right\|_{p_c \to p_c'} \lesssim \lambda^{\frac{d}{2}\left(\frac{1}{p_c}-\frac{1}{p_c'}\right)-1}. \tag{17.3}$$

We remark that the above change of variables is just saying

$$dE_{-\Delta}(\lambda) = \frac{1}{2}\lambda^{\frac{d}{2}-1} R^*_{\sqrt{\lambda}\mathbb{S}^{d-1}} R_{\sqrt{\lambda}\mathbb{S}^{d-1}}\, d\lambda\,.$$

One could have obtained (17.3) also from Stone's formula

$$\frac{1}{2}((E(\Lambda)f,f)+(E(\overline{\Lambda})f,f)) = \lim_{\varepsilon\searrow 0}\frac{1}{2\pi i}\int_\Lambda ([R(\lambda+i\varepsilon)-R(\lambda-i\varepsilon)]f,f)\, d\lambda\,,$$

or equivalently (in the weak sense)

$$\frac{dE_A(\lambda)}{d\lambda} = \frac{1}{2\pi i}\left((A-(\lambda+i0))^{-1}-(A-(\lambda-i0))^{-1}\right) = \frac{1}{\pi}\mathrm{Im}\left((A-(\lambda+i0))^{-1}\right)\,,$$

(recall also that $\mathrm{Im}(F(E+i\varepsilon))\, dE \rightharpoonup d\mu(E)$ where $F(z)=\int(\lambda-z)^{-1}\, d\mu(\lambda)$ denotes the Borel transformation of the (spectral) measure μ) and the "uniform" (in $\mathrm{Im}(z)$) resolvent bound of Kenig–Ruiz–Sogge [120, Theorem 2.3], i.e.,

$$\sup_{\mathrm{Im}(z)\in(0,1)} \|(-\Delta-z)^{-1}\|_{p\to p'} \lesssim_{d,p} |z|^{-(d+2)/2+d/p} = |z|^{\frac{d}{2}\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \tag{17.4}$$

for all $2d/(d+2)\le p\le 2(d+1)/(d+3)=p_c$ which, in turn, is obtained via complex interpolation between the L^2 boundedness of

$$T_\zeta = \frac{\mathrm{e}^{\zeta^2}}{\Gamma(d/2+\zeta)}(-\Delta-z)^\zeta$$

for $\mathrm{Re}(\zeta)=0$ and the $L^1\to L^\infty$ boundedness for $\mathrm{Re}(\zeta)\in[-(d+1)/2,-d/2]$. In turn, the latter follows from the explicit expression of the Fourier transform of the symbol of T_ζ.

Remark 17.1. The analogous estimate

$$\|((-\Delta)^{s/2}-z)^{-1}\|_{L^p(\mathbb{R}^d)\to L^{p'}(\mathbb{R}^d)} \lesssim |z|^{\frac{d}{s}\left(\frac{1}{p}-\frac{1}{p'}\right)-1}, \quad s\in\left[\frac{2d}{d+1},d\right), \quad p\in\left[\frac{2d}{d+s},\frac{2(d+1)}{d+3}\right] \tag{17.5}$$

was proved by Cuenin [53] (in fact also for more general operators including Dirac) and Huang et al [112].

We will now upgrade (17.3)[19] using the observation

$$dE_A(\lambda) = 2^{2k}\left(1+\frac{A}{\lambda}\right)^{-2k} dE_A(\lambda)$$

and the estimate

$$\begin{aligned}\|(1-\Delta/\lambda)^{-k}\|_{p\to q} &= \frac{1}{\Gamma(k)}\|\int_0^\infty \mathrm{e}^{-t(1-\Delta/\lambda)}t^{k-1}\,dt\|_{p\to q} \leq \frac{1}{\Gamma(k)}\int_0^\infty \mathrm{e}^{-t}(t/\lambda)^{-\frac{d}{2}\left(\frac{1}{p}-\frac{1}{q}\right)}t^{k-1}\,dt \\ &\lesssim \lambda^{\frac{d}{2}\left(\frac{1}{p}-\frac{1}{q}\right)}\end{aligned}$$

for $2k > d(1/p-1/q)$. Indeed, this estimate and (17.3) yield

$$\begin{aligned}\left\|\frac{dE_{-\Delta}(\lambda)}{d\lambda}\right\|_{1\to\infty} &= 2^{2k}\left\|(1-\Delta/\lambda)^{-k}\frac{dE_{-\Delta}(\lambda)}{d\lambda}(1-\Delta/\lambda)^{-k}\right\|_{1\to\infty} \\ &\lesssim_k \|(1-\Delta/\lambda)^{-k}\|_{1\to p_c}\|(1-\Delta/\lambda)^{-k}\|_{p_c'\to\infty}\left\|\frac{dE_{-\Delta}(\lambda)}{d\lambda}\right\|_{p_c\to p_c'} \lesssim \lambda^{d/2-1}\,.\end{aligned}$$

Thus, by interpolation, (17.3) can be upgraded to

$$\left\|\frac{dE_{-\Delta}(\lambda)}{d\lambda}\right\|_{p\to p'} \lesssim \lambda^{\frac{d}{2}\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \tag{17.6}$$

for all $1 \leq p \leq p_c$.

Let us finally mention that (17.3) respectively (17.6) should have always be more precisely written as

$$\|\mathbf{1}_{[\lambda,\lambda+1]}(-\Delta)\|_{p\to p'} \lesssim \lambda^{\frac{d}{2}\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \tag{17.7}$$

which just follows from setting $F(k) = \mathbf{1}_{[\lambda,\lambda+1]}(k)$ in (17.2). Formulae (17.3) respectively (17.6) correspond to the choice $F(k) = \delta(k-\lambda)$.

17.2. **Restriction theorems and multiplier theorems for Schrödinger operators.** By perturbation theory, the resolvent estimate (17.4) and the spectral projection estimate (17.6) can be further upgraded to treat $-\Delta+V$ for $0 \leq V \in L^{d/2}\cap L^{d/2+\varepsilon}$ for some $\varepsilon > 0$, see Ionescu–Schlag [114] (for uniform resolvent estimates which imply spectral measure estimates by Stone) or Huang et al [112]. (In fact, the non-negativity of V is only used to prove the bound on $\|(1-\Delta/\lambda)^{-k}\|_{p\to q}$ when one applies Trotter's formula, i.e., ultimately to prove the $L^1 \to L^\infty$ bound on $dE_\Delta(\lambda)$; neither the resolvent bound, nor the $L^{p_c} \to L^{p_c'}$ bound on $dE_\Delta(\lambda)$ use that V is non-negative.) It is for this very reason that estimates like (17.1) and (17.1) are sometimes called *Tomas–Stein estimates* as well, see, e.g., [155, p. 3073-3074].

For further generalizations of the above theme, we refer to the works by Guillarmou et al [99], Sikora et al [154, 155], and Chen et al [45, 46].

17.3. **Distorted Fourier transform.** In the following we consider Schrödinger operators of the form

$$H = P_0(D) + V(x,D) \quad \text{in } L^2(\mathbb{R}^d)$$

where P_0 is real and simply characteristic (see Hörmander [109, Definition 14.3.1]), $\sigma_{pp}(P_0) = \{0\}$, and $V(x,D)$ is a symmetric short range perturbation of P_0 in the sense of Hörmander [109,

[19]This is no upgrade as the restriction estimate in (17.3) already holds for all $p \in [1,p_c]$.

Definition 14.4.1]. Recall the Agmon–Hörmander spaces B and B^* (see, e.g., Hörmander [109, Section 14.1]) and let

$$Z(P_0) := \{\lambda \in \mathbb{R} :\ P_0(\xi) = \lambda \text{ and } dP_0(\xi) = 0 \text{ for some } \xi \in \mathbb{R}^d\} \quad \text{and}$$
$$S_\lambda := \{\xi \in \mathbb{R}^d : P_0(\xi) = \lambda\} .$$

Recall that

$$\begin{aligned}\int \mathbf{1}_\Omega(\lambda)(dE_\lambda^{(0)} f, f) &= \pm \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \int_{\mathbb{R}} \mathbf{1}_\Omega(\lambda) \mathrm{Im}(R_0(\lambda \pm i\varepsilon) f, f)\, d\lambda \\ &= \int_{\mathbb{R}} d\lambda\, \mathbf{1}_\Omega(\lambda) \int_{S_\lambda} |\hat{f}(\xi)|^2\, d\sigma_{S_\lambda}(\xi) , \quad f \in L^2 .\end{aligned}$$

Recall the resolvent formula $R(\lambda \pm i0) f = R_0(\lambda \pm i0) f_{\lambda \pm i0}$ where $f_z = (1 + V R_0(z))^{-1} f$ is a continuous function of $z \in \mathbb{C}^{\pm} \setminus (\sigma_{pp}(H) \cup Z(P_0))$ with values in B. Thus, we have

$$\int \mathbf{1}_\Omega(\lambda)(dE_\lambda^{(V)} f, f) = \int_{\mathbb{R}} d\lambda\, \mathbf{1}_\Omega(\lambda) \int_{S_\lambda} |\hat{f}_{\lambda \pm i0}(\xi)|^2\, d\sigma_{S_\lambda}(\xi) , \quad f \in B ,$$

whenever $\Omega \cap (\sigma_{pp}(H) \cup Z(P_0)) = \emptyset$. This motivates

Definition 17.2. If $f \in B$, then the L^2 functions defined by

$$\begin{aligned}(\mathcal{F}_\pm f)(\xi) &= \mathcal{F}[(1 + V R_0(\lambda \pm i0))^{-1} f](\xi) , \quad \xi \in S_\lambda \\ &= \mathcal{F}[(1 - V R(\lambda \pm i0)) f](\xi)\end{aligned} \tag{17.8}$$

almost everywhere in S_λ are called *distorted Fourier transforms of f*.

We recall the following properties of solutions of scattering states. Let $B_{P_0}^* = \{u :\ P_0^{(\alpha)} u \in B^* \text{ for every } \alpha\}$.

Lemma 17.3 (Hörmander [109, Lemma 14.6.6])**.** *If $u \in B_{P_0}^*$, $\lambda \notin Z(P_0)$, and $(P_0(D) + V - \lambda) u = 0$, then u is given by the solution of the* Lippmann–Schwinger *equation*

$$u = u_\pm - R_0(\lambda \mp i0) V u \tag{17.9}$$
$$= (1 - R(\lambda \mp i0) V) u_\pm , \tag{17.10}$$

where

$$\hat{u}_\pm = v_\pm \delta(P_0 - \lambda) = v_\pm d\sigma_{S_\lambda}(\xi) , \quad v_\pm \in L^2(S_\lambda, d\Sigma_{S_\lambda})$$

and

$$\int_{S_\lambda} (|v_+|^2 - |v_-|^2)\, d\sigma_{S_\lambda}(\xi) = 0 \tag{17.11}$$

where $d\sigma_{S_\lambda}(\xi) = |\nabla P_0(\xi)|^{-1} d\Sigma_{S_\lambda}(\xi)$ and $d\Sigma_{S_\lambda}(\xi)$ is the euclidean surface measure on S_λ. Moreover, if $\lambda \notin (Z(P_0) \cup \sigma_{pp}(P_0 + V))$, then

$$(\mathcal{F}_+ f, \hat{u}_+) = (\mathcal{F}_-, \hat{u}_-) = (f, u) , \quad \text{if } f \in B . \tag{17.12}$$

Let us also recall

Theorem 17.4 (Hörmander [109, Lemma 14.6.4 and Theorem 14.6.5])**.** *$\mathcal{F}_\pm : E^c L^2(\mathbb{R}^d) \to \widehat{L^2(\mathbb{R}^d)}$ is an isometric operator, which vanishes on $E^{pp} L^2(\mathbb{R}^d)$, with*

$$\|E^c f\|_2^2 = \int_{\mathbb{R}^d} |\mathcal{F}_\pm f(\xi)|^2\, d\xi .$$

Moreover, the intertwining property

$$\mathcal{F}_\pm \mathrm{e}^{itH} = \mathrm{e}^{itP_0(\xi)} \mathcal{F}_\pm$$

holds for all $t \in \mathbb{R}$. In particular, the restriction of H to $E^c L^2$ is absolutely continuous (since P_0 has purely absolutely continuous spectrum).

Moreover, $\mathcal{F}_\pm : E^c L^2(\mathbb{R}^d) \to \widehat{L^2(\mathbb{R}^d)}$ is actually unitary, i.e., the restriction of H to $E^c L^2$ is unitarily equivalent to P_0, i.e., $\sigma_c(H) = \sigma_{ac}(H) = \sigma(P_0)$. In particular, for $f \in E^c(L^2(\mathbb{R}^d))$, we have

$$(\mathcal{F}_\pm H f)(\xi) = P_0(\xi)(\mathcal{F}_\pm f)(\xi)\,, \quad \textit{i.e.,} \quad (Hf)(x) = (\mathcal{F}_\pm^* P_0(\cdot)\mathcal{F}_\pm f)(x)\,.$$

In particular, it follows that

$$\mathcal{F}_\pm^* \mathcal{F}_\pm = E^c \quad \text{and} \quad \mathcal{F}_\pm \mathcal{F}_\pm^* = \mathbf{1}_{\widehat{L^2}}\,.$$

The distorted Fourier transform (17.8) can be conveniently represented using the solutions $\varphi_{\xi(\lambda)}(x)$ (for $\xi(\lambda) \in S_\lambda$) of the Lippmann–Schwinger equation (17.9). In fact, we have (see also Ikebe [113] and Yafaev [202, Sections 6.6-6.8])

$$(\mathcal{F}_\pm f)(\xi) = \langle \varphi_\xi, f\rangle\,, \quad \xi \in \bigcup_{\lambda \in \sigma_{ac}(H)} S_\lambda \tag{17.13}$$

$$(\mathcal{F}_\pm^* g)(x) = \int_{\mathbb{R}^d} \varphi_\xi(x) g(\xi)\, d\xi = \int_{\sigma_{ac}(H)} d\lambda \int_{S_\lambda} d\sigma_{S_\lambda}(\xi)\ \varphi_\xi(x) g(\xi)\,. \tag{17.14}$$

Moreover, we have the following expansion theorem (see also Ikebe [113, Theorem 5])

$$f = \sum_{\lambda \in \sigma_{pp}(H)} |\psi_\lambda\rangle\langle\psi_\lambda, f\rangle + \int_{\mathbb{R}^d} |\varphi_\xi\rangle\langle\varphi_\xi, f\rangle\, d\xi \tag{17.15}$$

where $\{\psi_\lambda\}_{\lambda \in \sigma_{pp}(H)}$ denote the L^2-normalized eigenfunctions of H, i.e., $H\psi_\lambda = \lambda\psi_\lambda$. Moreover,

$$Hf = \sum_{\lambda \in \sigma_{pp}(H)} \lambda |\psi_\lambda\rangle\langle\psi_\lambda, f\rangle + \int_{\mathbb{R}^d} P_0(\xi) |\varphi_\xi\rangle\langle\varphi_\xi, f\rangle\, d\xi\,. \tag{17.16}$$

The above results motivate in particular the following definition of the distorted Fourier restriction and extension operators

$$(F_{S_\lambda} f)(\xi) = \langle \varphi_\xi, f\rangle = (\mathcal{F}_\pm f)(\xi)\,, \quad \xi \in S_\lambda \tag{17.17}$$

$$(F_{S_\lambda}^* g)(x) = \int_{S_\lambda} d\sigma_{S_\lambda}(\xi)\ \varphi_{\xi(\lambda)}(x) g(\xi) \tag{17.18}$$

which are defined with respect to the canonical measure $d\sigma_{S_\lambda}$. In particular, we have for any $\Lambda \subseteq \sigma_{ac}(H)$,

$$E_H(\Lambda) = \int_{P_0^{-1}(\Lambda)} |\varphi_\xi\rangle\langle\varphi_\xi|\, d\xi = \int_\Lambda d\lambda \int_{S_\lambda} d\sigma_{S_\lambda}(\xi)\ |\varphi_{\xi(\lambda)}\rangle\langle\varphi_{\xi(\lambda)}| = \int_\Lambda d\lambda\ F_{S_\lambda}^* F_{S_\lambda}$$

in a suitable weak sense and in particular, for $\lambda \in \sigma_{ac}(H)$,

$$\frac{dE_H(\lambda)}{d\lambda} = \int_{S_\lambda} d\sigma_{S_\lambda}(\xi)\ |\varphi_{\xi(\lambda)}\rangle\langle\varphi_{\xi(\lambda)}| = F_{S_\lambda}^* F_{S_\lambda}\,.$$

17.4. **Eigenfunction estimates for $F(-\Delta)$.** The theme in the first subsection can clearly be generalized. We are picking up the discussion from Remark 2.1.

Suppose, we are given a continuous function $a : \mathbb{R}^d \to [0,\infty)$ with

$$\nabla a(\xi) \neq 0 \quad \text{for } \xi \in a^{-1}(\Lambda)\,, \Lambda \subseteq \mathbb{R}\,.$$

Then we can define the Fourier multiplier $H_0 = \mathcal{F}^* A \mathcal{F}$, where A is multiplication by the symbol $a(\xi)$ and $X \subseteq \mathbb{R}$ is some Borel set. It is well known that its spectral projection is given by

$$E(X) = \mathcal{F}^* \mathbf{1}_{\{a^{-1}(X)\}} \mathcal{F}\,.$$

Now consider the "cospheres" associated to a,

$$S_\lambda := \{\xi \in \mathbb{R}^d : a(\xi) = \lambda\}$$

with the associated Lebesgue surface measure $d\sigma_\lambda(\xi)$. We may then define the *canonical measure associated to a* by

$$d\Sigma_\lambda(\xi) = \frac{d\sigma_\lambda(\xi)}{|\nabla a(\xi)|}\,,$$

which is, however, not intrinsic to S_λ. (See also Strichartz [172, p. 705].) In particular, the elementary volume $d\xi$ in $\mathbb{R}^d$ satisfies

$$d\xi = d\lambda d\Sigma_\lambda(\xi)\,.$$

Thus, by the above discussion, we can write the spectral projection E as

$$\langle\psi, E(X)\psi\rangle = \int_{\mathbb{R}^d} \mathbf{1}_{\{\xi\in\mathbb{R}^d : a(\xi)\in X\}}(\xi)|\hat\psi(\xi)|^2\,d\xi = \int_{a^{-1}(X)} |\hat\psi(\xi)|^2\,d\xi = \int_X d\lambda \int_{S_\lambda} |\hat\psi(\xi)|^2\,d\Sigma_\lambda(\xi)\,. \tag{17.19}$$

Thus, $a(D)$ has absolutely continuous spectrum and the spectral projection-valued measure is given by $dE(\lambda) = F_{S_\lambda}^* F_{S_\lambda} d\lambda$, where F_{S_λ} denotes the Fourier restriction operator associated to the measure $d\Sigma_\lambda$ (see below) and we have

$$\frac{dE(\lambda)}{d\lambda} f(x) = \int_{S_\lambda} \mathrm{e}^{2\pi i x\cdot\xi} \hat f(\xi)\,d\Sigma_\lambda(\xi)\,.$$

In particular, for a given measurable function $F : [0,\infty) \to \mathbb{R}$, we have

$$\langle\psi, F(H_0)\psi\rangle = \int_{\mathbb{R}_+} d\lambda\; F(\lambda) \int_{S_\lambda} |\hat\psi(\xi)|^2\,d\Sigma_\lambda(\xi)\,.$$

It follows from (17.19) that the space $E(\Lambda)\mathcal{H}$, in which H_0 becomes diagonal, is given by the direct integral

$$E(\Lambda)\mathcal{H} \leftrightarrow \int_\Lambda^\oplus L^2(S_\lambda)\,d\lambda\,,$$

where S_λ is endowed with the measure $d\Sigma_\lambda$. A vector $f \in E(\Lambda)\mathcal{H}$ is mapped in this direct integral into an element $\tilde f(\lambda)$ which, for every fixed $\lambda \in \Lambda$, is the restriction of $\hat f$ on S_λ.

Let us now denote the Fourier restriction and extension operators on S_λ by F_{S_λ} and $F_{S_\lambda}^*$ and abbreviate F_S, respectively F_S^* if $\lambda = 0$. In particular,

$$F_{S_\lambda}^* \varphi(x) = \int_{S_\lambda} \mathrm{e}^{2\pi i x\cdot\xi} \varphi(\xi)\,d\Sigma_\lambda(\xi)\,.$$

Now, if $S_{\lambda=0}$ has non-vanishing curvature and a is sufficiently smooth, it follows again by the Tomas–Stein theorem that the associated spectral projection

$$\frac{dE(\lambda)}{d\lambda}(\lambda = 0) = F_S^* F_S$$

satisfies

$$\left\|\frac{dE(\lambda)}{d\lambda}(\lambda = 0)\right\|_{p_c\to p_c'} \lesssim 1\,,$$

where $p_c = 2(d+1)/(d+3)$. One may now ask how these estimates behave, when one varies λ. Clearly, the bounds depend heavily on the restriction estimates for S_λ, and thus, it is inevitable to control the behavior of the surface measure $d\Sigma_\lambda$ as λ varies.

Since the spectral measure dE is absolutely continuous, we have

$$dE(\lambda) = \frac{dE(\lambda)}{d\lambda}\,d\lambda = F_{S_\lambda}^* F_{S_\lambda}\,d\lambda = \left(\widehat{d\Sigma_{S_\lambda}} *\right)\,d\lambda\,,$$

i.e., it suffices to control $\|\widehat{d\Sigma_{S_\lambda}}\|_r$ for $1/r = 1+1/p'-1/p$ (and possibly $p = p_c = 2(d+1)/(d+3)$). For a given diffeomorphism $\psi(\lambda) : S \to S_\lambda$, the Radon–Nikodým derivative is given by

$$\tau(\lambda,\zeta) := \frac{d\Sigma_{S_\lambda}(\psi(\lambda)\zeta)}{d\Sigma_S(\zeta)} = \exp\left(\int_0^\lambda d\mu\ (\mathrm{div}\ j)(\psi(\mu)\zeta)\right), \quad \zeta \in S\,,$$

where $j(\xi) = |\nabla P(\xi)|^{-2}\nabla P(\xi)$, see, e.g., Yafaev [202, Lemma 2.1.9]. Thus, we have

$$\widehat{d\Sigma_{S_\lambda}}(x) = \int_{S_\lambda} \mathrm{e}^{2\pi i x\cdot\xi}\, d\Sigma_{S_\lambda}(\xi) = \int_S \mathrm{e}^{2\pi i x\cdot\psi(\lambda)\zeta}\tau(\lambda,\zeta)\, d\Sigma_S(\zeta)\,,$$

and therefore

$$\|\widehat{d\Sigma_{S_\lambda}}\|_r \leq \sup_{\zeta\in S}\tau(\lambda,\zeta)\|\widehat{d\Sigma_S}(\tilde{\psi}(\lambda)\cdot)\|_r\,,$$

where $\tilde{\psi}(\lambda) : \mathbb{R}^d \to \mathbb{R}^d$ is defined by

$$\langle\tilde{\psi}(\lambda)x,\zeta\rangle = \langle x,\psi(\lambda)\zeta\rangle\,.$$

Example 17.5. (1) For $P(\xi) = \xi^2$ (i.e., $P(D) = -\Delta$), we take $S = \mathbb{S}^{d-1}$, and $S_\lambda = \{\xi \in \mathbb{R}^d : \xi^2 = \lambda\} = \sqrt{\lambda}\mathbb{S}^{d-1}$. We reparameterize $\lambda = 1 \pm \rho$ for $\rho > 0$, i.e., $S_{\lambda(\rho)} = \{\xi \in \mathbb{R}^d : |\xi^2 - 1| = \rho\}$. (For the "inner" surface, we restrict to $\rho < 1$ of course.) Thus, for our "new symbol" $P(\xi) = \xi^2 - 1$ defining $S_{\rho(\lambda)}$, we have $j(\xi) = \xi/(2|\xi|^2)$ and div $j(\xi) = (d-2)/(2|\xi|^2)$. We will now construct a C^1-diffeomorphism $\psi : S \to S_\lambda$ with $\psi(0)\zeta = \zeta$ and $\psi(\rho)\zeta = \sqrt{1\pm\rho}\zeta$. For $t \geq 0$, we define $\psi(t)\zeta = \sqrt{1\pm t}\zeta$. Thus,

$$\tau(\rho,\zeta) = \exp\left(\int_0^\rho d\mu \frac{d-2}{2(1\pm\mu)}\right) = (1\pm\rho)^{(d-2)/2} = \lambda^{(d-2)/2}$$

and we obtain

$$\|\widehat{d\Sigma_{S_\lambda}}\|_r = \lambda^{(d-2)/2}\cdot\lambda^{-d/(2r)}\|\widehat{d\Sigma_S}\|_r \lesssim \lambda^{d/2-1-d/2+d(1/p-1/p')/2} = \lambda^{d(1/p-1/p')/2-1}\,,$$

thereby recovering (17.3).

(2) For $P(\xi) = |\xi|$, i.e., $P(D) = \sqrt{-\Delta}$, we take $S = \mathbb{S}^{d-1}$ and $S_\lambda = \{\xi \in \mathbb{R}^d : |\xi| = \lambda\} = \lambda\mathbb{S}^{d-1}$. The situation is pretty clear since

$$\widehat{d\Sigma_{S_\lambda}}(x) = \int_{S_\lambda} \mathrm{e}^{2\pi i x\cdot\xi}\, d\Sigma_{S_\lambda}(\xi) = \lambda^{d-1}\int_S \mathrm{e}^{2\pi i x\cdot(\lambda\xi)}\, d\Sigma_S(\xi) = \lambda^{d-1}\widehat{d\Sigma_S}(\lambda x)\,.$$

Thus, we immediately obtain

$$\|\widehat{d\Sigma_{S_\lambda}}\|_r = \lambda^{d-1}\cdot\lambda^{-d/r}\|\widehat{d\Sigma_S}\|_r \lesssim \lambda^{d-1-d+d(1/p-1/p')} = \lambda^{d(1/p-1/p')-1}\,,$$

thereby recovering (17.1). In principle, one could go through the above steps and explicitly construct a diffeomorphism $\psi(\lambda) : S \to S_\lambda$ and compute the Radon–Nikdým derivative; but since the situation here is so simple, we refrain from doing so.

17.5. **An application of the observation of Frank and Sabin.** We follow [87, Section 4]. Our goal is to prove uniform Sobolev estimates and limiting absorption principles (LAPs) for Schrödinger operators in Schatten ideals. We begin with the former which is an extension to the uniform Sobolev estimate by Kenig–Ruiz–Sogge [120, Theorem 2.3].

Theorem 17.6 (Uniform Sobolev estimate in Schatten spaces)**.** *Let $d \geq 2$ and assume that*

$$\begin{cases} q \in [4/3, 3/2] & \text{if } d = 2\,,\\ q \in [d/2, (d+1)/2] & \text{if } d \geq 3\,.\end{cases}$$

Then for all $z \in \mathbb{C} \setminus [0,\infty)$, we have the estimates

$$\|W_1(-\Delta - z)^{-1}W_2\|_{\mathcal{S}^{(d-1)q/(d-q)}(L^2(\mathbb{R}^d))} \lesssim |z|^{-1+d/(2q)}\|W_1\|_{L^{2q}}\|W_2\|_{L^{2q}} \tag{17.20}$$

and, for $\gamma \geq 1/2$, $\delta(z) := \operatorname{dist}(z,[0,\infty))$, and all $z \in \mathbb{C} \setminus [0,\infty)$,

$$\|W_1(-\Delta - z)^{-1}W_2\|_{\mathcal{S}^{2(\gamma+d/2)}(L^2(\mathbb{R}^d))} \lesssim \delta(z)^{-1+\frac{(d+1)/2}{2(\gamma+d/2)}}|z|^{-\frac{1}{2(\gamma+d/2)}}\|W_1\|_{L^{2(\gamma+d/2)}}\|W_2\|_{L^{2(\gamma+d/2)}} \,. \tag{17.21}$$

Proof. We begin with the proof of (17.20). By scaling it suffices to consider $z \in \mathbb{C} \setminus \{1\}$ with $|z| = 1$. For such z we will prove the bounds

$$\|W_1^{-it}(-\Delta - z)^{it}W_2^{-it}\|_{L^2\to L^2} \leq \|W_1\|_\infty\|W_2\|_\infty \,, \quad t \in \mathbb{R} \tag{17.22}$$

and

$$\|W_1^{a-it}(-\Delta - z)^{-a+it}W_2^{a-it}\|_{\mathcal{S}^2} \leq M_{d,a}\mathrm{e}^{C_{d,a}t^2}\|W_1\|^a_{L^{\frac{4ad}{d-1+2a}}}\|W_2\|^a_{L^{\frac{4ad}{d-1+2a}}} \,, \quad t \in \mathbb{R} \tag{17.23}$$

where a is an arbitrary parameter satisfying $1 \leq a \leq 3/2$ if $d = 2$ and $d/2 \leq a \leq (d+1)/2$ if $d \geq 3$. Obviously, these estimates imply

$$\begin{aligned}
&\|W_1(-\Delta - z)^{it}W_2\|_{L^2\to L^2} \leq \|W_1\|_\infty\|W_2\|_\infty \\
&\|W_1(-\Delta - z)^{-a+it}W_2\|_{\mathcal{S}^2} \leq M_{d,a}\mathrm{e}^{C_{d,a}t^2}\|W_1\|_{L^{\frac{4d}{d-1+2a}}}\|W_2\|_{L^{\frac{4d}{d-1+2a}}}
\end{aligned}$$

for all $t \in \mathbb{R}$. Thus, complex interpolation for Schatten ideals[20] (cf. Simon [156, Theorem 2.9]) applied to the family $W_1(-\Delta - z)^{-\zeta}W_2$ then gives

$$\|W_1(-\Delta - z)^{-1}W_2\|_{\mathcal{S}^{2a}} \lesssim \|W_1\|_{L^{\frac{4ad}{d-1+2a}}}\|W_2\|_{L^{\frac{4ad}{d-1+2a}}} \,.$$

Up to the change of variables $a = q(d-1)/(2(d-q))$ this is just the claimed estimate. In fact, if $d = 2,3$ and $a = 1$, then (17.23) is already the desired bound and complex interpolation is not necessary.

So let us prove (17.22) and (17.23). The former estimate is an immediate consequence of Plancherel. For the latter, we will estimate $|(-\Delta - z)^{-a+it}(x-y)|$ and apply either

- the Hardy–Littlewood–Sobolev inequality if it is bounded by a constant times $|x-y|^{-\zeta}$ for some $\zeta \in (0,d)$ or
- Hölder's inequality if it is uniformly bounded in $|x-y|$. (This is the case when $4d/(d-1+2a) = 2$, i.e., $a = (d+1)/2$. In this case the L^2 norms of $|W_1|^a$ and $|W_2|^a$ are taken on the right side of (17.23), as expected.)

To that end recall [120, Formulae (2.21), (2.23), (2.25)], i.e.,

$$(-\Delta - z)^\lambda(x-y) = \frac{2^{\lambda+1}}{(2\pi)^{d/2}\Gamma(-\lambda)}\left(\frac{z}{|x-y|^2}\right)^{\frac{d/2+\lambda}{2}} K_{\frac{d}{2}+\lambda}\left(\sqrt{z}\,|x-y|\right)$$

and, with $\nu \in \mathbb{C}$,

$$\begin{aligned}
&|\mathrm{e}^{\nu^2}\nu K_\nu(w)| \leq C|w|^{-|\mathrm{Re}(\nu)|} && \text{for } |w| \leq 1,\ \mathrm{Re}(w) > 0\,, \\
&|K_\nu(w)| \leq C_{\mathrm{Re}(\nu)}\mathrm{e}^{-\mathrm{Re}(w)}|w|^{-1/2} && \text{for } |w| \geq 1,\ \mathrm{Re}(w) > 0,\ \mathrm{Re}(\nu) \geq 0\,.
\end{aligned}$$

[20] Note that although (17.23) deteriorates super-exponentially, it is still sub-double-exponential in t, so Stein interpolation is indeed applicable.

Setting $\lambda = -a + it$, $\nu = d/2 + \lambda = d/2 - a + it$, we have $\operatorname{Re}\nu \in [0, 1/2]$ for $a \in [d/2, (d+1)/2]$. Thus, for $w = \sqrt{z}\,|x-y|$ with $z \neq 1$ but $|z| = 1$, i.e., $|w| = |x-y|$, we can estimate in this case

$$\begin{aligned} |K_\nu(w)| &\lesssim_{a,d} \mathrm{e}^{c_{d,a}t^2} \left[|w|^{-|\mathrm{Re}(\nu)|} \wedge |w|^{-1/2} \left(1 \wedge \mathrm{Re}(w)^{-N}\right) \right], \quad && \mathrm{Re}(w) > 0,\ N \in \mathbb{N} \\ &\lesssim_{a,d} \mathrm{e}^{c_{d,a}t^2} |x-y|^{-1/2}, && |w| = |x-y|,\ |\mathrm{Re}(\nu)| \leq 1/2 . \end{aligned}$$

Combining the previous estimates therefore gives

$$\begin{aligned} |(-\Delta - z)^{-a+it}(x-y)| &\lesssim_{a,d} \frac{2^{1-a}}{(2\pi)^{d/2}|\Gamma(a-it)|} \left(\frac{|z|}{|x-y|^2}\right)^{\frac{d/2-a}{2}} \mathrm{e}^{c_{d,a}t^2}|x-y|^{-1/2} \\ &\lesssim_{a,d} \mathrm{e}^{c_{d,a}t^2}|x-y|^{a-\frac{d+1}{2}} . \end{aligned}$$

Thus, by the Hardy–Littlewood–Sobolev inequality, we obtain

$$\|W_1^{a-it}(-\Delta - z)^{-a+it}W_2^{a-it}\|_{\mathcal{S}^2}^2 \leq M_{d,a}\mathrm{e}^{C_{d,a}t^2}\|W_1\|_{\frac{4ad}{d-1+2a}}^{2a}\|W_2\|_{\frac{4ad}{d-1+2a}}^{2a} \quad \text{if } a \in \left[\frac{d-1}{2}, \frac{d+1}{2}\right]$$

for all $t \in \mathbb{R}$. This is precisely (17.23) and concludes the proof of (17.20).

The second estimate (17.21) follows from complex interpolation between (17.20) for $\gamma = 1/2$ respectively $q = (d+1)/2$, i.e.,

$$\|W_1(-\Delta - z)^{-1}W_2\|_{\mathcal{S}^{d+1}} \lesssim |z|^{-1/(d+1)}\|W_1\|_{L^{d+1}}\|W_2\|_{L^{d+1}} \tag{17.24}$$

and the trivial bound for $\gamma = \infty$, i.e.,

$$\|W_1(-\Delta - z)^{-1}W_2\| \leq \delta(z)^{-1}\|W_1\|_\infty\|W_2\|_\infty . \tag{17.25}$$

This concludes the proof. □

We will now use and upgrade arguments of Ionescu–Schlag [114] to obtain a LAP in Schatten spaces for $V \in L^q$. As in their arguments, a crucial ingredient is a deep result of Koch and Tataru [125, Theorem 3] about absence of embedded eigenvalues for such potentials.

Theorem 17.7 (LAP for $V \in L^q$ in Schatten spaces)**.** *Let $d \geq 2$ and assume that $V \in L^q(\mathbb{R}^d : \mathbb{R})$ with*

$$\begin{cases} q \in (1, 3/2] & \text{if } d = 2, \\ q \in [d/2, (d+1)/2] & \text{if } d \geq 3. \end{cases}$$

Define $\alpha_q := 2 \vee (d-1)q/(d-q)$. Then

(1) $V^{1/2}(-\Delta + V - z)^{-1}|V|^{1/2} \in \mathcal{S}^{\alpha_q}(L^2(\mathbb{R}^d))$ for every $z \in \mathbb{C} \setminus [0,\infty)$.

(2) the mapping $\mathbb{C} \setminus [0,\infty) \ni z \mapsto V^{1/2}(-\Delta + V - z)^{-1}|V|^{1/2} \in \mathcal{S}^{\alpha_q}$ is analytic and extends continuously to $(0,\infty)$ (with possibly different boundary values from above and below).

(3) under the additional assumption $q > d/2$, there is a constant $C_{d,q}$ (independent of V) such that for $|z|^{-1+d/(2q)}\|V\|_{L^q} \leq C_{d,q}$, one has

$$\|V^{1/2}(-\Delta + V - z)^{-1}|V|^{1/2}\|_{\mathcal{S}^{\alpha_q}} \leq 2C_{d,q}|z|^{-1+d/(2q)}\|V\|_{L^q} . \tag{17.26}$$

If $q = d/2$ and $d \geq 3$, the bound (17.26) holds provided $|z| \geq C(V)$ for some constant $C(V)$ only depending on V.

The proof of this theorem relies on detailed information of the Birman–Schwinger operator $V^{1/2}(-\Delta - z)^{-1}|V|^{1/2}$.

Lemma 17.8. *Let $d \geq 2$ and assume $V \in L^q(\mathbb{R}^d)$ where q satisfies the assumptions in Theorem 17.7. Let $I \subseteq (0,\infty)$ be a compact interval. Then*

(1) the family
$$A(z) := V^{1/2}(-\Delta - z)^{-1}|V|^{1/2} \in \mathcal{S}^{\alpha_q}(L^2(\mathbb{R}^d))$$
is analytic on the half strips $S_\pm := \{z \in \mathbb{C} : \operatorname{Re}(z) \in \mathring{I},\ \pm\operatorname{Im}(z) > 0\}$.

(2) On each $S_\pm$*, the family* $A(z)$ *is continuous up to* $\overline{S_\pm}$ *and we denote by* $V^{1/2}(-\Delta - \lambda \pm i0)^{-1}|V|^{1/2}$ *its extensions at* $\lambda > 0$.

(3) For all $z \in \overline{S_\pm}$ *we have the estimate*
$$\|A(z)\|_{\mathcal{S}^{\alpha_q}} \leq C|z|^{-1+d/(2q)}\|V\|_{L^q}\,, \tag{17.27}$$
where C *is the implicit constant in* (17.20) *which is, in particular, independent of* I.

(4) for all $z \in \overline{S_\pm}$ *the operator* $1 + A(z)$ *is invertible and the map* $S_\pm \ni z \mapsto (1 + A(z))^{-1}$ *is an analytic family of bounded operators on* $L^2(\mathbb{R}^d)$ *which is continuous on* $\overline{S_\pm}$.

It is precisely this lemma which relies on the absence of embedded eigenvalues [125, Theorem 3]. The proof of Theorem 17.7 is then a simple combination of the uniform Sobolev inequalities of Theorem 17.6 and this lemma (together with the resolvent identity).

Proof of Lemma 17.8. (1) The family $\mathbb{C} \setminus [0,\infty) \ni z \mapsto V^{1/2}(-\Delta - z)^{-1}|V|^{1/2}$ is indeed analytic as can be seen by invoking the resolvent formula. We obtain for any $z, z_0 \in \mathbb{C} \setminus [0,\infty)$,
$$\begin{aligned}V^{1/2}(-\Delta - z)^{-1}|V|^{1/2} &- \sum_{n=0}^{N}(z - z_0)^n V^{1/2}(-\Delta - z_0)^{-n-1}|V|^{1/2}\\ &= V^{1/2}(-\Delta - z)^{-1}(z - z_0)^{N+1}(-\Delta - z_0)^{-N-1}|V|^{1/2}\,.\end{aligned}$$
By the Seiler–Simon inequality and the constraint $q \geq d/2$, the right side is bounded in $\mathcal{S}^{\alpha_q}$ norm by
$$\begin{aligned}&\|V^{1/2}(-\Delta - z)^{-1}(-\Delta - z_0)^{-N-1}|V|^{1/2}\|_{\mathcal{S}^{\alpha_q}}\\ &\quad\leq \||V|^{1/2}(-\Delta - z_0)^{-1}\|^2_{\mathcal{S}^{2\alpha_q}}\|(-\Delta - z_0)^{-1}\|^{N-1}\|(-\Delta - z)^{-1}\| \leq C^N\|V\|_q\end{aligned}$$
and hence vanishes as $N \to \infty$ if $|z - z_0|$ is small enough (such that $|z - z_0| \ll C^{-1}$ for instance). This shows that the entire series converges in $\mathcal{S}^{\alpha_q}$ with a nonzero convergence radius and thereby the asserted analyticity of $A(z)$ in $\mathcal{S}^{\alpha_q}$.

(2) Next, we notice that one can rely on the arguments and results of Ionescu–Schlag [114] as V is an admissible potential in their sense, see also [87, p. 1676]. In particular, [114, Lemma 4.1 b)] yields that for each $\lambda > 0$ there exists an operator $(-\Delta - \lambda \pm i0)^{-1} \in \mathcal{B}(L^{2q(q+1)} \to L^{2q(q-1)})$, i.e.,
$$\|(-\Delta - \lambda \pm i0)^{-1}\|_{L^{2q(q+1)} \to L^{2q(q-1)}} \leq C_I \quad \text{for any } \lambda \in \mathring{I}$$
such that $z \mapsto A(z)$ can be extended as a continuous family on the strips $\overline{S_\pm}$ in weak operator topology, i.e., there are sequences $\mathring{I} \ni \lambda_n \to \lambda$ and $\varepsilon_n \to 0$ such that
$$\lim_{n\to\infty}((-\Delta - \lambda_n \pm i\varepsilon_n)^{-1}f, \varphi) = ((-\Delta - \lambda \pm i0)^{-1}f, \varphi)\,, \quad f \in L^{2q/(q+1)},\ \varphi \in \mathcal{S}(\mathbb{R}^d)\,.$$
We will now show that this family is indeed continuous in $\mathcal{S}^{\alpha_q}$. To that end let $z \in \overline{S_\pm}$ and $(z_n) \subseteq S_\pm$ such that $z_n \to z$. Since the Schatten spaces are Banach, so in particular complete, it suffices to show that $A(z_n)$ is Cauchy in $\mathcal{S}^{\alpha_q}$ norm to show Schatten norm continuity of $A(z)$ up to the real axis. To that end, we decompose
$$\sqrt{V} = W_1 + W_1\,, \quad |V|^{1/2} = \tilde{W}_1 + \tilde{W}_2\,,$$
where $W_1, \tilde{W}_1$ are bounded, compactly supported functions and
$$\|W_2\|_{q/2} + \|\tilde{W}_2\|_{q/2} \leq \varepsilon\,.$$

Using the uniform Sobolev inequality (17.20), we then obtain

$$\|A(z_n) - A(z_m)\|_{\mathcal{S}^{\alpha_q}} \leq \|W_1((-\Delta - z_n)^{-1} - (-\Delta - z_m)^{-1})\tilde{W}_1\|_{\mathcal{S}^{\alpha_q}} + C\varepsilon\,.$$

The first term is easily bounded using the classic LAP in trace ideals for potentials that are short-range in pointwise sense, cf. Yafaev [202, Proposition VII.1.22]. (See also [202, Proposition VI.2.1] for Hölder continuity of the Birman–Schwinger in operator norm.) Anyway, that proposition asserts that the family $z \mapsto W_1(-\Delta - z)^{-1}\tilde{W}_1$ is analytic in $S_\pm$ and continuous on $\overline{S}_\pm$ in $\mathcal{S}^{\alpha_q}$ topology. In particular, it implies for n, m large enough

$$\|W_1((-\Delta - z_n)^{-1} - (-\Delta - z_m)^{-1})\tilde{W}_1\|_{\mathcal{S}^{\alpha_q}} \leq \varepsilon$$

for any given ε. Thus, $(A(z_n))_n$ is Cauchy in $\mathcal{S}^{\alpha_q}$ and hence $z \mapsto A(z) \in \mathcal{S}^{\alpha_q}$ is continuous up to the real line, i.e., the boundary of $S_\pm$. Let us repeat that this implies in particular that $V^{1/2}(-\Delta - \lambda \pm i0)^{-1}|V|^{1/2} \in \mathcal{S}^{\alpha_q}$ for all $\lambda > 0$ and that the asserted estimate (17.27) continuous carries over to the real axis.

(3) We apply analytic Fredholm theory (cf. Yafaev [201, Lemma I.8.1 and Theorems I.8.2-3]) to the family (of compact operators) $A(z)$ in the strips $\overline{S_\pm}$ and infer that $z \mapsto (1 + A(z))^{-1}$ is a meromorphic family of operators on $S_\pm$ with poles at those points z where $-1 \in \sigma(A(z))$. Moreover, this family is continuous up to the real axis, except at those points $\lambda \in I$ where $-1 \in \sigma(A(\lambda))$.

This almost finishes the proof, as we are left to show that no such points $z \in S_\pm$ exist such that $-1 \in \sigma(A(z))$. Recall that our potential V is assumed to be real-valued.

Case $\mathrm{Im}(z) \neq 0$**.** This follows from a simple argument similar to the one at the beginning of the proof of [114, Lemma 4.6]. We present the argument for the sake of completeness. We sandwich $1 + A(z)$ from the left with $|V|^{1/2}f$ and from the right with $V^{1/2}f$. Then, also the the imaginary part of

$$(f, Vf) + (f, V(-\Delta - \lambda \pm i0)^{-1}f) = 0$$

vanishes. Since the first summand is zero for real-valued V, so must be the second one, i.e.,

$$0 = \mathrm{Im}\int_{\mathbb{R}^d} |\widehat{Vf}|^2(\xi^2 - \lambda \pm i\varepsilon)^{-1}\,d\xi = \varepsilon \int_{\mathbb{R}^d} |\widehat{Vf}|^2\left[\varepsilon^2 + (\xi^2 - \lambda)^2\right]^{-1}$$

Since $\varepsilon > 0$, the integral must vanish, so the integrand is zero almost everywhere, i.e., $f \equiv 0$. So $1 + A(z)$ is invertible for all z with $\mathrm{Im}(z) > 0$ if V is real-valued.

Case $z > 0$**.** This is where the result on absence of embedded eigenvalue of Koch and Tataru comes in. So suppose there are $\lambda > 0$, a sign $\pm$, and $f \in L^2(\mathbb{R}^d)$ such that (for $(-\Delta - \lambda \pm i0)^{-1} \equiv R_0(\lambda)$)

$$V^{1/2}R_0(\lambda)|V|^{1/2}f = -f\,.$$

We will now show $f \equiv 0$. So let us define $g := R_0(\lambda)|V|^{1/2}f$. Since $f \in L^2$ and $V \in L^q$, we have $|V|^{1/2}f \in L^{2q/(q+1)}$ and by the classic uniform Sobolev inequality [120, Theorem 2.3] $g \in L^{2q/(q-1)}(\mathbb{R}^d)$. Moreover, $Vg \in L^{2q/(q+1)}$ and the above equation reads

$$R_0(\lambda)Vg = -g\,.$$

By the integrability properties of g and Vg, we can rewrite the equation as the well-defined Schrödinger equation $(-\Delta + V)g = zg$ in the sense of distributions on $\mathbb{R}^d$. Since $g \in L^{2q/(q-1)}$ and $Vg \in L^{2q/(q+1)}$, we have $g \in H^{2q/(q+1)}_{\mathrm{loc}} \subseteq H^1_{\mathrm{loc}}$. Once we show that $g \in L^2$ (or $|x|^{-1/2+\varepsilon}g \in L^2$ for some $\varepsilon > 0$), we can apply [125, Theorem 3] and conclude $g \equiv 0$ and therefore also $f = -V^{1/2}R_0(\lambda)|V|^{1/2} = -\sqrt{V}g \equiv 0$ and therefore $-1 \notin \sigma(A(z))$.

So we are left to show $g \in L^2$. Since $Vg \in L^{2q/(q+1)}$, we have $Vg \in X$ where X denotes the Banach space defined in the introduction of Ionescu–Schlag [114] (that plays a similar role than the Agmon–Hörmander spaces). By [114, Lemma 4.1 a,b)], we know $R_0(\lambda) : X \to X^*$ boundedly. Thus, $g = -R_0(\lambda)Vg \in X^*$. Using $R_0(\lambda)Vg = -g$ and [114, Lemma 4.4], we obtain

$$\|(1+|x|^2)^M g\|_{X^*} < \infty\,, \quad M \geq 0\,.$$

Writing $g =< x >^{-2M}< x >^{2M} g$ and recalling $X^* \subseteq L^{2q/(q-1)}$, we see $g \in L^2$. This finally concludes the proof of Lemma 17.8. □

We are now ready to prove Theorem 17.7. It basically uses the resolvent identity to upgrade the results of Lemma 17.8 on $V^{1/2}(-\Delta - z)^{-1}|V|^{1/2}$ to $V^{1/2}(-\Delta + V - z)^{-1}|V|^{1/2}$.

Proof of Theorem 17.7. We rewrite the operator of interest as

$$V^{1/2}(-\Delta+V-z)^{-1}|V|^{1/2} = \frac{1}{1+V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}} V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}\,. \tag{17.28}$$

By Lemma 17.8, we know that the maps

$$z \mapsto \frac{1}{1+V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}} \in \mathcal{B}(L^2(\mathbb{R}^d))\,, \quad z \mapsto V^{1/2}(-\Delta-z)^{-1}|V|^{1/2} \in \mathcal{S}^{\alpha_q}(L^2(\mathbb{R}^d))$$

are analytic on $\mathbb{C} \setminus [0,\infty)$ and extend continuously to $(0,\infty)$ with possibly different boundary values from above and below. This settles (1) and (2).

Thus, we are left to prove the uniform Schatten bound (17.26). Indeed, for $q > d/2$ and $z \in \mathbb{C} \setminus [0,\infty)$ such that $C|z|^{-1+d/(2q)}\|V\|_{L^q} \leq 1/2$ we obtain (by (17.28), the Schatten bound for the Birman–Schwinger operator in (17.27), and the uniform resolvent estimate for Schatten spaces in Theorem 17.6)

$$\begin{aligned}
&\|V^{1/2}(-\Delta+V-z)^{-1}|V|^{1/2}\|_{\mathcal{S}^{\alpha_q}} \\
&\leq \left\| \left(1+V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}\right)^{-1} \right\| \cdot \|V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}\|_{\mathcal{S}^{\alpha_q}} \\
&\leq \left(\sum_{n\geq 0} \|V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}\|^n \right) \cdot C|z|^{-1+d/(2q)}\|V\|_{L^q} \\
&\leq 2C|z|^{-1+d/(2q)}\|V\|_{L^q}\,, \qquad \text{for } C|z|^{-1+d/(2q)}\|V\|_q \leq 1/2\,.
\end{aligned}$$

Finally, let $q = d/2$ and $d \geq 3$. Similarly as in the proof of Lemma 17.8 we decompose $V^{1/2} = W_1 + W_2$ and $|V|^{1/2} = \tilde{W}_1 + \tilde{W}_2$ with $W_1, \tilde{W}_1 \in C_c$ bounded and $W_2, \tilde{W}_2 \in L^{q/2}$ with $L^{q/2}$ norm $< \varepsilon$. Then, again by the uniform Sobolev estimate in Schatten spaces (Theorem 17.6),

$$\|V^{1/2}(-\Delta-z)^{-1}|V|^{1/2}\| \leq \|W_1(-\Delta-z)^{-1}\tilde{W}_1\| + C\varepsilon\,, \quad z \in \mathbb{C} \setminus [0,\infty)\,.$$

But since $W_1, \tilde{W}_1$ also belong to $L^{q/2}$ for any $q > d/2$, we can apply our previous result and infer

$$\|W_1(-\Delta-z)^{-1}\tilde{W}_1\| \to 0 \quad \text{as } |z| \to \infty\,.$$

Thus, there is a $C(V)$ such that for all $|z| \geq C(V)$ we obtain the same bound (17.26). This concludes the proof of Theorem 17.7. □

18. Connection between Fourier restriction and eigenvalue estimates

We are mainly motivated by [83] who used the Birman–Schwinger principle and the Kenig–Ruiz–Sogge estimate to obtain bounds on every eigenvalue of $-\Delta - V$ in $L^2(\mathbb{R}^d)$ with complex-valued V. In [57], the Stein–Tomas estimate for Schatten ideals, Theorem 4.15, was used to analyze eigenvalue sums of $|\Delta+1| - V$ with real-valued V. (It transpires from the approach that the kinetic energy can be much more general.)

18.1. **Definition of resolvent for large class of potentials.** To analyze eigenvalues of $-\Delta + V$, it is helpful to understand how the resolvent

$$R_V(\lambda) = (-\Delta + V - \lambda^2)^{-1}, \quad \lambda \in \mathbb{C} \tag{18.1}$$

is defined.

18.1.1. *Compactly supported potentials.*

Theorem 18.1 ([70, Theorem 3.8])**.** *Let $d \geq 3$ be odd and $V \in L^\infty_{\mathrm{comp}}(\mathbb{R}^d)$. If $\mathrm{Im}(\lambda) > 0$, then $R_V(\lambda) : L^2 \to L^2$ is a meromorphic family of operators with finitely many poles. In fact, it extends to a meromorphic family of operators $R_V(\lambda) : L^2_{\mathrm{comp}} \to L^2_{\mathrm{loc}}$ for all $\lambda \in \mathbb{C}$.*

Definition 18.2. Let $V \in L^\infty_{\mathrm{comp}}(\mathbb{R}^d)$. Then the poles of the meromorphic continuation $R_V(\lambda) : L^2_{\mathrm{comp}} \to L^2_{\mathrm{loc}}$ with $\lambda \in \mathbb{C}$ are called *scattering resonances*. See [70, 55]

For recent *effective bounds on the number of resonances* $\#\{\text{resonances}\lambda : |\lambda| < r\}$, see [55].

18.1.2. *L^p-potentials.* When analyzing real-valued potentials and $\lambda \in \mathbb{R}$, we may assume $V \in L^q$ with $q \in [1, (d+1)/2]$. If not, we have to sacrifice the end-point $(d+1)/2$ and assume there is $\gamma > 0$ such that $V \in \mathrm{e}^{-\gamma|\cdot|} L^q$ with $q \in [1, (d+1)/2)$ or $V \in \mathrm{e}^{-\gamma|\cdot|} L^{(d+1)/2,1}$ at most. We use the notation

$$v_0 := \|V\|_{(d+1)/2} \quad \text{and} \quad v_\gamma : -\|\mathrm{e}^{2\gamma|\cdot|} V\|_{(d+1)/2,1} \tag{18.2}$$

with the Lorentz norm

$$\|V\|_{p,q} := p^{\frac{1}{q}} \left(\int_0^\infty \frac{d\alpha}{\alpha} \alpha^q |\{x \in \mathbb{R}^d : |V(x)| > \alpha\}|^{q/p} \right)^{1/q}, \quad p, q \in (0, \infty) \tag{18.3}$$

$$\|V\|_{p,\infty} := \sup_{\alpha > 0} \alpha |\{x \in \mathbb{R}^d : |V(x)| > \alpha\}|^{1/p}. \tag{18.4}$$

Define the Birman–Schwinger operator

$$BS(\lambda) := |V|^{1/2} R_0(\lambda) V^{1/2}. \tag{18.5}$$

If V is bounded, then by iterating the second resolvent identity,

$$R_V(\lambda) = R_0(\lambda) - R_0(\lambda) V^{1/2} (1 + BS(\lambda))^{-1} |V|^{1/2} R_0(\lambda). \tag{18.6}$$

This formula is valid for $\mathrm{Im}\lambda \gg 1$ since, by the Stein–Tomas theorem

$$\|BS(\lambda)\| \lesssim |\lambda^{-\frac{2}{d+1}} \|V\|_{(d+1)/2}, \quad \mathrm{Im}\lambda > 0. \tag{18.7}$$

If V is unbounded, one can *use* (18.6) *as a definition* of $R_V(\lambda)$ for $\mathrm{Im}\lambda \gg 1$. This idea goes back to Kato [116], but see also [93, 94] for abstract results in the non-self-adjoint setting.

Recall the Stein–Tomas theorem for Schatten ideals (Theorem 4.15) saying

$$\|BS(\lambda)\|_{d+1} \lesssim \lambda^{-\frac{2}{d+1}} \|V\|_{(d+1)/2}, \quad \mathrm{Im}(\lambda) \geq 0. \tag{18.8}$$

On the other hand, Frank–Laptev–Safronov [86] considered numbers of eigenvalues of Schrödinger operators with complex V and extended this to $\mathrm{Im}\lambda < 0$. They showed

$$\|BS(\lambda)\|_{d+1} \lesssim \lambda^{-\frac{2}{d+1}} \|\mathrm{e}^{2\beta_d (\mathrm{Im}\lambda)_- /(d+1)\,|\cdot|} V\|_{(d+1)/2}, \quad \lambda) \in \mathbb{C}, \quad \text{for some } \beta_d \geq \frac{2(\mathrm{e}^{(d+1)/2} - 1)}{\mathrm{e} - 1}. \tag{18.9}$$

In particular, $BS(\lambda)$ is compact whenever $\beta_d (\mathrm{Im}\lambda)_- /(d+1) < \gamma$ and $v_\gamma < \infty$. By the meromorphic Fredholm theorem (see, e.g., [70, Theorem C.8]), the definition (18.6) then defines a meromorphic continuation of R_V to the region $\mathrm{Im}\lambda > -(d+1)\gamma/\beta_d$; in partiular, scattering resonances are then poles of the operator-valued function $\lambda \mapsto (1 + BS(\lambda))^{-1}$. Moreover, it turns out [55, Corollary 3.2] that, for $(\mathrm{Im}\lambda)_- < \gamma$, $BS(\lambda) \in \mathcal{S}^{(d+1),\infty}$, the weak Schatten class consisting of compact operators whose singular values s_n obey $\sup_{n\in\mathbb{N}} n^{1/(d+1)} s_n < \infty$.

Theorem 18.3 ([55, Proposition 2.1])**.** *If $d \geq 3$ is odd and $V \in L^{(d+1)/2,1}$, then $R_V(\lambda)$ has a meromorphic continuation to* $\mathrm{Im}(\lambda) > -\gamma$.

18.1.3. *Pointwise decaying potentials.* See [55, Section 2].

18.2. **Alternative proof of Stein–Tomas inequality in Schatten spaces.** We provide an alternative proof for the Stein–Tomas theorem in Schatten spaces (Theorem 4.15) which does not rely on complex interpolation. Such proof can be helpful to investigate the isospectral operator $\mathcal{E}^*V\mathcal{E}$ and allowes to keep possible valuable oscillation of V.

Theorem 18.4. *Let $V \in L^{(d+1)/2,1}$. Then $\|\mathcal{E}^*V\mathcal{E}\|_{\mathcal{S}^{d+1,\infty}} \lesssim \|V\|_{\frac{d+1}{2},1}$.*

Proof. We perform a dyadic horizontal decomposition

$$V(x) = \sum_{j\geq 0} V_j \quad \text{with} \quad V_j = V\mathbf{1}_{H_j \leq |V| \leq H_{j+1}} \tag{18.10}$$

where

$$H_j = \inf\{t > 0 : \, |\{|V| > t\}| \leq 2^{j-1}. \tag{18.11}$$

This is called horizontal decomposition, because $|\mathrm{supp}(V_j)| \sim 2^j$. In particular,

$$\|V\|_{L^{q,r}} \sim \|H_j 2^{j/q}\|_{\ell^r_j(\mathbb{Z}_+)}. \tag{18.12}$$

Then, by the triangle inequality and the fact that $|V_j| \sim H_j$,

$$\|\mathcal{E}^*V\mathcal{E}\|_{\mathcal{S}^{d+1,\infty}} \lesssim \sum_{j\in\mathbb{Z}_+} H_j \|\mathcal{E}^*\mathbf{1}_{\mathrm{supp}(V_j)}\mathcal{E}\|_{\mathcal{S}^{d+1,\infty}}. \tag{18.13}$$

In Theorem 18.5 below, we show

$$s_n(\mathcal{E}^*\mathbf{1}_\Omega\mathcal{E}) \lesssim n^{-\frac{1}{d+1}} |\Omega|^{\frac{2}{d+1}}, \quad \Omega \text{ measurable}, \tag{18.14}$$

which, together with the definition of weak Schatten spaces[21], yields

$$\|\mathcal{E}^*\mathbf{1}_{\mathrm{supp}(V_j)}\mathcal{E}\|_{\mathcal{S}^{d+1,\infty}} \lesssim |\mathrm{supp}(V_j)|^{\frac{2}{d+1}} \sim 2^{\frac{2j}{d+1}}. \tag{18.15}$$

This concludes the proof. □

Theorem 18.5. *Let $\Omega \subseteq \mathbb{R}^d$ be measurable and $s_k(\mathcal{E}^*\mathbf{1}_\Omega\mathcal{E})$ be the k-th singular value of $\mathcal{E}^*\mathbf{1}_\Omega\mathcal{E}$. Then*

$$s_k(\mathcal{E}^*\mathbf{1}_\Omega\mathcal{E}) = s_k(\mathbf{1}_\Omega\mathcal{E}\mathcal{E}^*\mathbf{1}_\Omega) \lesssim k^{-\frac{1}{d+1}} |\Omega|^{\frac{2}{d+1}}. \tag{18.16}$$

The following proof is not efficient for oscillating potentials, since we entirely work with $\mathbf{1}_\Omega\mathcal{E}\mathcal{E}^*\mathbf{1}_\Omega$.

Proof. Let $\Omega = \bigcup_{\alpha\in\rho\mathbb{Z}^d} Q_\alpha$ be a disjoint partition of Ω into cubes $Q_\alpha = \rho\alpha + [0,\rho]^d$ with side length ρ centered at $\alpha \in \rho\mathbb{Z}^d$. Let $\chi_\alpha := \mathbf{1}_{Q_\alpha}$. We split

$$T := \mathbf{1}_\Omega\mathcal{E}\mathcal{E}^*\mathbf{1}_\Omega = T_1 + T_2 \tag{18.17}$$

with

$$T_1(s) := \sum_{|\alpha-\beta|>s} \chi_\alpha\mathcal{E}\mathcal{E}^*\chi_\beta \quad \text{and} \quad T_2(s) := \sum_{|\alpha-\beta|\leq s} \chi_\alpha\mathcal{E}\mathcal{E}^*\chi_\beta. \tag{18.18}$$

---

[21] Recall that the singular values $s_n(T)$ of $T \in \mathcal{S}^{p,\infty}$ obey $\sup_{n\in\mathbb{N}} n^{1/p} s_n(T) \sim \|T\|_{p,\infty} < \infty$.

We now use Fan's inequality $s_{n+m+1}(A+B) \leq s_{n+1}(A) + s_{m+1}(B)$ for all $n, m \geq 0$ and obtain

$$s_n(T) = s_n(T_1 + T_2) \leq s_n(T_1) + \|T_2\|. \tag{18.19}$$

We begin with estimating $\|T_2\|$. To that end, we use Plancherel's theorem and regard $S = \{(\xi', \varphi(\xi') : \xi' \in \mathcal{D} \subseteq \mathbb{R}^{d-1}\}$ as the graph of a function $\varphi : \mathcal{D} \to \mathbb{R}$ with $\mathcal{D} \subseteq \mathbb{R}^{d-1}$. Then

$$(\mathcal{E}g)(x) = \int_{\mathbb{R}^{d-1}} e(x_d \varphi(\xi') + x' \cdot \xi') g(\xi', \varphi(\xi')) \mathbf{1}_{\mathcal{D}}(\xi') \, d\xi'. \tag{18.20}$$

Writing $F_{x_d}(\xi') = e(x_d \varphi(\xi')) g(\xi', \varphi(\xi')) \mathbf{1}_{\mathcal{D}}(\xi')$ and using Plancherel in $\mathbb{R}^{d-1}$ yields

$$\begin{aligned} \|\chi_\alpha \mathcal{E} g\|_{L^2(\mathbb{R}^d)}^2 &\leq \int_{\alpha_d}^{\alpha_d+\rho} dx_d \int_{\mathbb{R}^{d-1}} dx' \Big| \int_{\mathbb{R}^{d-1}} d\xi' \, F_{x_d}(\xi') \Big|^2 \\ &= \int_{\alpha_d}^{\alpha_d+\rho} dx_d \int_{\mathbb{R}^{d-1}} d\xi' |F_{x_d}(\xi')|^2 \\ &= \int_{\alpha_d}^{\alpha_d+\rho} dx_d \int_{\mathbb{R}^{d-1}} d\xi' |g(\xi', \varphi(\xi')) \mathbf{1}_{\mathcal{D}}(\xi')|^2 = \rho \|g\|_{L^2(S)}^2. \end{aligned} \tag{18.21}$$

This shows, by Young's inequality for sums,

$$\begin{aligned} \|T_2(s)\| &\leq \sup_{g_1, g_2 \in L^2(\mathbb{R}^d), \|g_1\|, \|g_2\| \leq 1} \sum_{|\alpha-\beta| \leq s} |\langle \mathcal{E}^* \chi_\alpha g_1, \mathcal{E}^* \chi_\beta g_2 \rangle| \mathbf{1}_{|\alpha-\beta| \leq s} \\ &\leq \sup_{\|g\|_2 \leq 1} \sum_{\alpha \in \rho\mathbb{Z}^d} \|\mathcal{E}^* \chi_\alpha g\|_2^2 \cdot \sum_{|\beta| \leq s} 1 \\ &\lesssim \rho s^d \sup_{\|g\|_2 \leq 1} \sum_{\alpha \in \rho\mathbb{Z}^d} \|\chi_\alpha g\|_2^2 \leq \rho s^d. \end{aligned} \tag{18.22}$$

To estimate $s_n(T_1)$, we use

$$s_n(T_1) \leq \|T_1\|_{p,\infty} n^{-1/p} = n^{-1/p} \sup_{\lambda > 0} \lambda (\#\{k : s_k(T_1) \geq \lambda\})^{1/p}. \tag{18.23}$$

We take $p = 2$ and estimate the right side using Markov's inequality,

$$\#\{k : s_k(T_1) \geq \lambda\} \leq k^{-2} \|T_1\|_2^2. \tag{18.24}$$

The integral kernel of T_1 is

$$T_1(x, y) = \sum_{|\alpha-\beta| > s} \chi_\alpha(x) \chi_\beta(y) (\mathcal{E}\mathcal{E}^*)(x - y). \tag{18.25}$$

Since $|(\mathcal{E}\mathcal{E}^*)(z)| \sim \langle z \rangle^{-(d-1)/2}$, we have

$$\begin{aligned} \|T_1\|_2^2 &= \int_{\mathbb{R}^{2d}} dx \, dy \left| \sum_{|\alpha-\beta| > s} \chi_\alpha(x) \chi_\beta(y) (\mathcal{E}\mathcal{E}^*)(x - y) \right|^2 \\ &\lesssim (\rho s)^{-(d-1)} \int_{\mathbb{R}^{2d}} dx \, dy \sum_{\alpha, \beta} \chi_\alpha(x) \chi_\beta(y) \\ &= (\rho s)^{-(d-1)} \sum_{\alpha, \beta} |Q_\alpha| |Q_\beta| \lesssim (\rho s)^{-(d-1)} |\Omega|^2. \end{aligned} \tag{18.26}$$

Here we used

$$\left[\sum_{\alpha,\beta} \chi_\alpha(x)\chi_\beta(y)\right]^2 = \sum_{\alpha,\beta} \chi_\alpha(x)\chi_\beta(y) \tag{18.27}$$

by the disjointness of the cubes Q_α, and that the summations over α and β only run over such $\alpha, \beta \in \rho\mathbb{Z}^d$ for which $Q_\alpha \cap \Omega \neq \emptyset$ and $Q_\beta \cap \Omega \neq \emptyset$. Thus, combining (18.23)–(18.26) shows

$$s_n(T_1) \lesssim n^{-1/2}\,(s\rho)^{-(d-1)/2}|\Omega|. \tag{18.28}$$

Combining this with (18.19)–(18.22) shows

$$s_n(T) \lesssim n^{-1/2}\,(s\rho)^{-(d-1)/2}|\Omega| + s^d\rho. \tag{18.29}$$

Optimizing over ρ, with optimizer ρ_* satisfying $\rho_*^{\frac{d+1}{2}} = c(s)|\Omega|n^{-1/2}$ shows

$$s_n(T) \lesssim c(s)|\Omega|^{2/(d+1)}n^{-1/(d+1)} \tag{18.30}$$

as desired. This concludes the proof. □

18.3. **Estimates for singular values.** Recall Stone's formula saying that $dE(k^2) = \pi^{-1}\mathrm{Im}((-\Delta - k^2 \pm i0)^{-1}) = c_d k^{\frac{d-2}{2}}\mathcal{E}(k)\mathcal{E}^*(k)$ for all $k \in \mathbb{R}$ with the scaled extension operator

$$\begin{aligned} &\mathcal{E}(k) : L^1(\mathbb{S}^{d-1}) \to L^\infty_{\mathrm{comp}}(\mathbb{R}^d) \\ &\mathcal{E}(k)g(x) = \int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi i k x\cdot\xi} g(\xi)\,d\omega(\xi), \quad x \in \mathbb{R}^d,\ k \in \mathbb{R}, \end{aligned} \tag{18.31}$$

where $d\omega(\xi)$ denotes the induced Lebesgue measure on $\mathbb{S}^{d-1}$. Clearly, $\mathcal{E}(k)$ can be analytically continued to $k \in \mathbb{C}$. In this case, we have, from Stone

$$R_0(\lambda) - R_0(-\lambda) = c_d\lambda^{d-2}\mathcal{E}(\lambda)\mathcal{E}(\overline{\lambda}), \quad \mathrm{Im}\lambda < 0. \tag{18.32}$$

We record the scaled Stein–Tomas and Agmon–Hörmander estimates,

$$\|\mathcal{E}(\lambda)g\|_{L^{\frac{2(d+1)}{d-1}}(\mathbb{R}^d)} \lesssim |\lambda|^{-\frac{d(d-1)}{2(d+1)}}\|g\|_{L^2(\mathbb{S}^{d-1})} \tag{18.33}$$

$$\|\mathcal{E}(\lambda)g\|_{L^2(B(R))} \lesssim R^{\frac12}|\lambda|^{-\frac{d-1}{2}}\|g\|_{L^2(\mathbb{S}^{d-1})}. \tag{18.34}$$

In fact, in combination with the Stein–Tomas estimate for Schatten spaces, one has

$$\||V^{1/2}\mathcal{E}(\lambda)\mathcal{E}(\lambda)^*V^{1/2}\|_{d+1} \lesssim |\lambda|^{-\frac{d(d-1)}{d+1}}\|V\|_{(d+1)/2}, \quad \mathrm{Im}\lambda \geq 0. \tag{18.35}$$

19. Cuenin's toolbox

We collect some standard tools of Cuenin.

19.1. **Phragmén–Lindelöf.** We follow [149, Section 5] and start with the Helmholtz equation

$$(\Delta + k^2)u = f \tag{19.1}$$

whose solution is

$$u(x) = \int \Phi(x-y)f(y)\,dy \tag{19.2}$$

where

$$\Phi(x-y) = c_d k^{\frac{d-2}{2}}\frac{H^{(1)}_{(d-2)/2}(k|x-y|)}{|x-y|^{(d-2)/2}}, \quad \text{where } c_d = \frac{1}{2i(2\pi)^{(d-2)/2}} \tag{19.3}$$

where $H^{(1)}_\lambda = J_\lambda + iY_\lambda$ is the Hankel function of the first kind, which coincides, up to innocent prefactors, with the modified Bessel function $K_\lambda(\pm ix)$ with imaginary argument. (Recall that the modulus of Hankel functions decays like $|x|^{-1/2}$, thus, $|\Phi(x)| \lesssim \langle x\rangle^{-(d-1)/2}$ for $k = 1$.) Clearly,

$$\hat{\Phi}(\xi) = \left(-\xi^2 + k^2 + i0\right)^{-1} . \tag{19.4}$$

This distribution can be expressed in terms of homogeneous distributions of degree -1, principal value (p.v.), and Dirac delta, see, e.g., [92, pp. 209–236]. We can obtain the expression from the one variable formula

$$\lim_{\varepsilon\searrow 0}(t + i\varepsilon)^{-1} = \text{p.v.}\frac{1}{t} + i\pi\delta \tag{19.5}$$

which can be extended to $\mathbb{R}^d$-function $t = H(\xi)$ (e.g., $H(\xi) = \xi^2$) as far as we can take locally H as a coordinate function in a local patch of a neighborhood in $\mathbb{R}^d$ at any point ξ_0 at which $H(\xi_0) = 0$. We have

Proposition 19.1. *Let $H : \mathbb{R}^d \to \mathbb{R}$ be such that $|\nabla H(\xi)| \neq 0$ at any point ξ where $H(\xi) = 0$. Then we can define the distribution limit*

$$(H(\xi) + i0)^{-1} = \lim_{\varepsilon\searrow 0}(H(\xi) + i\varepsilon)^{-1} , \tag{19.6}$$

and we have

$$(H(\xi) + i0)^{-1} = \text{p.v.}\frac{1}{H(\xi)} + i\pi\delta(H(\xi)) \tag{19.7}$$

in distributional sense.

In the above $\delta(H(\xi))$ is defined via

$$\delta(H)[\psi] := \int_{\mathbb{R}^d} \delta(H(\xi) - 0)\psi(\xi)\, d\xi = \int\limits_{\{H(\xi)=0\}} \psi(\xi)\, d\sigma_H(\xi) \tag{19.8}$$

where $d\sigma_H(\xi) = \frac{d\Sigma_H(\xi)}{|\nabla H(\xi)|}$ is the Leray measure with $d\Sigma_H(\xi)$ being the Lebesgue measure induced by $d\xi$ on the hypersurface $\{H(\xi) = 0\}$. We summarize this in

Lemma 19.2. *Let $H(\xi) = -\xi^2 + k^2$. Then*

$$(H(\xi) + i0)^{-1} = \text{p.v.}\frac{1}{H(\xi)} + \frac{i\pi}{2k} d\Sigma_{\{\xi^2=k^2\}} . \tag{19.9}$$

Thus,

$$(R_+(k^2)f)(x) = ((\Delta + k^2 + i0)^{-1}f)(x) = \text{p.v.} \int_{\mathbb{R}^d} \frac{\hat{f}(\xi)\mathrm{e}^{2\pi i x\cdot\xi}}{-\xi^2 + k^2} + \frac{i\pi}{2k}(d\sigma^{\vee}_{k\mathbb{S}^{d-1}} * f)(x) . \tag{19.10}$$

We begin with the well-known KRS bound.

Theorem 19.3 ([149, Theorem 5.2])**.** *Let $k > 0$ and $\frac{2}{d} \geq \frac{1}{p} - \frac{1}{p'} \geq \frac{2}{d+1}$ and $d \geq 3$ or $1 > \frac{1}{p} - \frac{1}{p'} \geq \frac{2}{3}$ and $d = 2$. Then*

$$\|R_+(k^2)f\|_{L^{p'}} \lesssim k^{-2+d(\frac{1}{p}-\frac{1}{p'})} . \tag{19.11}$$

The proof of the endpoint estimate was carried out earlier and used Stein's complex interpolation method. To treat $1/p - 1/p' > 2/(d + 1)$ it suffices to use real interpolation and the following estimates.

Lemma 19.4. *Let $\chi \in \mathcal{S}$ and denote $d\sigma_\varepsilon := f(\cdot)d\sigma(\cdot) * \varepsilon^{-d}\chi(\frac{\cdot}{\varepsilon})$, where $f \in L^\infty(\mathbb{S}^{d-1})$. Then*

$$\sup_\xi |d\sigma_\varepsilon(\xi)| \lesssim \|f\|_{L^\infty(\mathbb{S}^{d-1})}\varepsilon^{-1} \tag{19.12}$$

Proof. Without loss of generality suppose $f=1$. We first make the reduction to the case where χ is compactly supported. This part of the argument is called "Schwartz tail argument". Take a C_c^∞ partition of unity of $\mathbb{R}^d$ such that

$$\sum_{j\in\mathbb{N}_0}\psi_j(\xi)=1\,, \tag{19.13}$$

where ψ_0 is supported in $B_0(1)$ and $\psi_j(\xi)=\psi(\xi/2^j)$ for $j\geq 1$, where ψ is supported in the annulus $\{\xi\in\mathbb{R}^d:\,1/2\leq|\xi|\leq 2\}$. We adapt this partition to the resolution ε, i.e., we take

$$\sum_{j\in\mathbb{N}_0}\psi_j(\xi/\varepsilon)=1\,. \tag{19.14}$$

Now write

$$d\sigma_\varepsilon(\xi)=d\sigma(\cdot)*\varepsilon^{-d}\sum_{j\in\mathbb{N}_0}\psi_j(\frac{\cdot}{\varepsilon})\chi(\frac{\cdot}{\varepsilon})(\xi)\,. \tag{19.15}$$

Observe that for any $N\in\mathbb{N}$,

$$|\chi(\frac{\xi-\eta}{\varepsilon})|\lesssim_N\frac{1}{(1+2^j)^N}\,,\quad |\xi-\eta|\sim\varepsilon 2^j \tag{19.16}$$

and so

$$|d\sigma(\cdot)*\varepsilon^{-d}\psi_j(\frac{\cdot}{\varepsilon})\chi(\frac{\cdot}{\varepsilon})(\xi)|=\varepsilon^{-d}\left|\int_{\mathbb{S}^{d-1}}d\sigma(\eta)\psi_j(\frac{\xi-\eta}{\varepsilon})\chi(\frac{\xi-\eta}{\varepsilon})\right|\lesssim\frac{(2^j\varepsilon)^{d-1}}{(1+2^j)^N}\cdot\varepsilon^{-d}\,. \tag{19.17}$$

Taking N sufficiently large makes the j-summation convergent so that

$$d\sigma_\varepsilon(\xi)\lesssim\varepsilon^{-1}\,. \tag{19.18}$$

Since the $j=0$-term satisfies the estimate trivially, the proof is concluded. □

Lemma 19.5 ([149, Lemma 5.2])**.** *Let $\chi\in\mathcal{S}$ and*

$$R_\varepsilon(\xi)=\left(\frac{1}{-|\cdot|^2+k^2+i0}*\varepsilon^{-d}\chi(\frac{\cdot}{\varepsilon})\right)(\xi)\,. \tag{19.19}$$

Then

$$|R_\varepsilon(\xi)|\lesssim\varepsilon^{-1}\,, \tag{19.20}$$

and so in particular

$$|R_\varepsilon(\xi)|\lesssim\frac{1}{|-\xi^2+k^2|+\varepsilon}\,. \tag{19.21}$$

Moreover, for $\delta\in\mathbb{R}$, a Phragmén–Lindelöf argument combined with the previous estimates yields

$$\left(\frac{1}{-|\cdot|^2+k^2+i\delta}*\varepsilon^{-d}\chi(\frac{\cdot}{\varepsilon})\right)(\xi)\lesssim\frac{1}{|-\xi^2+k^2|+\varepsilon}\,. \tag{19.22}$$

First proof of Lemma 19.5. Explicit estimates in [149, Lemma 5.2]. Recall

$$\frac{1}{\Delta+1+i0}=\text{p.v.}\frac{1}{\Delta+1}+\frac{i\pi}{2}d\Sigma_{\mathbb{S}^{d-1}}\,. \tag{19.23}$$

By Lemma 19.4 it suffices to estimate the convolution with the first summand, i.e.,

$$P_\varepsilon(\xi):=\text{p.v.}\frac{1}{-\xi^2+1}*\varepsilon^{-d}\chi(\frac{\cdot}{\varepsilon})(\xi)\,. \tag{19.24}$$

Denoting $\chi_\varepsilon(\xi) = \varepsilon^{-d}\chi(\xi/\varepsilon)$, we shall estimate

$$\begin{aligned} P_\varepsilon(\xi) &= -\text{p.v.} \left(\int\limits_{1-\varepsilon<|\eta|<1+\varepsilon} + \int\limits_{|\eta|<1-\varepsilon} + \int\limits_{|\eta|>1+\varepsilon} \right) \chi_\varepsilon(\xi-\eta)\frac{1}{|\eta|^2-1}\,d\eta \\ &= I_1 + I_2 + I_3\,. \end{aligned} \tag{19.25}$$

The summands $I_2 + I_3 \lesssim \varepsilon^{-1}\|\chi\|_1$ are easily estimated. To estimate I_1 we write

$$\begin{aligned} I_1 &= \lim_{\delta\to 0} \int\limits_{\delta\le|1-|\eta||\le\varepsilon} \frac{\chi_\varepsilon(\xi-\eta)}{|\eta|^2-1}\,d\eta \\ &= \lim_{\delta\to 0}\left(\int_{1-\varepsilon}^{1-\delta} + \int_{1+\delta}^{1+\varepsilon}\right)\int_{\mathbb{S}^{d-1}} \chi_\varepsilon(\xi - r\theta)\frac{r^{d-1}}{(r+1)(r-1)}\,d\Sigma(\theta)\,. \end{aligned} \tag{19.26}$$

Changing $r = 2 - s$ in the second integral, we obtain

$$I_1 = \lim_{\delta\to 0}\int_{1-\varepsilon}^{1-\varepsilon} F(r,\xi)(r-1)^{-1}\,dr\,, \tag{19.27}$$

where

$$\begin{aligned} F(r,\xi) = &\int_{\mathbb{S}^{d-1}} \chi_\varepsilon(\xi-r\theta)\frac{r^{d-1}}{r+1}\,d\Sigma(\theta) \\ &- \int_{\mathbb{S}^{d-1}} \chi_\varepsilon(\xi-(2-r)\theta)\frac{(2-r)^{d-1}}{3-r}\,d\Sigma(\theta)\,. \end{aligned} \tag{19.28}$$

If we observe that $F(1,\xi) = 0$, we may write by the mean value theorem

$$\int_{1-\varepsilon}^{1-\delta} F(r,\xi)(r-1)^{-1}\,dr \le \varepsilon \sup_{1-\varepsilon\le r\le 1}\left|\frac{\partial F}{\partial r}(r,\xi)\right|\,. \tag{19.29}$$

The radial derivative of the first integral in the definition of $F(r,\xi)$ is given by

$$\frac{\partial}{\partial r}\left(\frac{r^{d-1}}{r+1}\right)\int_{\mathbb{S}^{d-1}} \chi_\varepsilon(\xi-r\theta)\,d\Sigma(\theta) + \frac{r^{d-1}}{r+1}\int_{\mathbb{S}^{d-1}} \theta\cdot\nabla\chi_\varepsilon(\xi-r\theta)\,d\Sigma(\theta)\,. \tag{19.30}$$

The second of these integrals can be written as

$$\varepsilon^{-1}\sum_{j=1}^{d}\frac{r^{d-1}}{r+1}\int_{\mathbb{S}^{d-1}}\theta_j\left(\frac{\partial}{\partial x_j}\chi\right)_\varepsilon(\xi-r\theta)\,d\Sigma(\theta)\,. \tag{19.31}$$

Thus, both integrals can be understood as mollifications with resolution ε of the measures $\theta_i\,d\Sigma(\theta)$, which, from Lemma 19.4, are bounded by $C(\chi)\varepsilon^{-1}$. This gives the desired estimate for the first integral in the definition of $F(r,\xi)$. The second integral can be treated in the same way. Thus,

$$\left|\frac{\partial F}{\partial r}(r,\xi)\right| \lesssim \varepsilon^{-2}\,, \tag{19.32}$$

and so we obtain the desired estimate $|I_1(\xi)| \lesssim \varepsilon^{-1}$. This concludes the proof. □

Second proof of Lemma 19.5. Alternatively (JC's arguments): by a partition of unity we may assume that $m(\xi) := (-\xi^2 + k^2 + i0)^{-1}$ is supported in a small conic neighborhood of the first coordinate axis. The implicit function theorem (see also (19.48) in the proof of Lemma 19.9 below) then allows us to reduce the proof to the bound

$$|\frac{1}{\xi_1 + i0} * \gamma_\varepsilon| \lesssim \varepsilon^{-1}\,, \tag{19.33}$$

where $\gamma_\varepsilon(\xi_1) = \varepsilon^{-1}\gamma(\xi_1/\varepsilon)$ is a function of one variable. By Riemann–Lebesgue and Hausdorff–Young,

$$|\frac{1}{\xi_1 + i0} * \gamma_\varepsilon| \lesssim \|\gamma_\varepsilon^\vee\|_1 \lesssim \varepsilon^{-1}\,, \tag{19.34}$$

where we also used that the Fourier transform of $(\xi_1 + i0)^{-1}$ is bounded, see also [111, Example 7.1.17]. The final estimate is a consequence of the estimate for $\delta = \pm 0$ and the Phragmén–Lindelöf principle, since better estimates are available for $\delta > 0$. Alternatively, one can appeal to the Malgrange preparation theorem [111, Theorem 7.5.5] and follow the proof of [13, Lemma 23], which we present later in Lemma 19.9. □

The KRS bound with zero imaginary part in Theorem 19.3 then implies

Theorem 19.6. *Let $z \in \mathbb{C}$ and $1/p - 1/p' \in [2/(d+1), 2/d]$ for $d \geq 3$ or $1/p - 1/p' \in [2/3, 1)$ for $d = 2$. Then for any $u \in C_c^\infty$, one has*

$$\|u\|_{p'} \lesssim |z|^{-1+d(\frac{1}{p}-\frac{1}{p'})}\|(\Delta + z)u\|_p\,. \tag{19.35}$$

The proof uses the following version of the Phragmén–Lindelöf maximum principle.

Proposition 19.7. *Let $F(z)$ be holomorphic in $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\} \equiv \mathbb{C}_+$ and continuous on the closure. Assume that $|F(z)| \leq L$ for $z \in \partial\mathbb{C}_+$ and that for any $\varepsilon > 0$ there is $C = C_\varepsilon$ such that $|F(z)| \leq C_\varepsilon e^{\varepsilon|z|}$ as $|z| \to \infty$, uniformly in the argument of z. Then $|F(z)| \leq L$ for any $z \in \mathbb{C}_+$.*

Proof of Theorem 19.6. Suppose $\hat{u}$ and $\hat{v}$ are compactly supported on $\mathbb{R}^d$ and consider in $\mathbb{C}_+$ the holomorphic function

$$F(z) = z^{1-\frac{d}{2}(\frac{1}{p}-\frac{1}{p'})}\langle v, (\Delta + z)^{-1}u\rangle = z^{1-\frac{d}{2}(\frac{1}{p}-\frac{1}{p'})}\langle \hat{v}, (-\xi^2 + z)^{-1}\hat{u}\rangle\,. \tag{19.36}$$

where we use the principal determination of $\log z$ and thereby of monomials $z^\alpha = e^{\alpha\log(z)}$ for $\alpha \in \mathbb{R}$. By the convergence of the distribution $(-\xi^2 + z + i\varepsilon)^{-1} \to (-\xi^2 + \mathrm{Re}(z) + i0)^{-1}$, we see that $F(z)$ is continuous on the closure of $\mathbb{C}_+$. By Theorem 19.3, we know $F(z) \leq C\|u\|_p\|v\|_p$ for $z \geq 0$, where C is independent of z. For $z \leq 0$, better, i.e., elliptic, estimates are available, so $F(z) \leq C\|u\|_p\|v\|_p$ for all $z \in \mathbb{R}$.

We now estimate $F(z)$ for $|z| \gg 1$. In particular, let $|z|$ be so large such that for all $\xi \in \mathrm{supp}(\hat{u})$, we have

$$(-|\xi|^2 + |z|)^{-1} \leq \frac{2}{|z|}\,. \tag{19.37}$$

Then we obtain

$$|F(z)| \leq C|z|^{1-\frac{d}{2}(\frac{1}{p}-\frac{1}{p'})} \cdot |z|^{-1}\int |\hat{u}(\xi)||\hat{v}(\xi)| \leq Ce^{\varepsilon|z|}\,, \quad \xi \in \mathrm{supp}(\hat{u})\,. \tag{19.38}$$

By the Phragmén–Lindelöf principle, we obtain

$$|F(z)| \leq C\|u\|_p\|v\|_p \Leftrightarrow |\langle v, (\Delta + z)^{-1}u\rangle| \leq C|z|^{-1+\frac{d}{2}(\frac{1}{p}-\frac{1}{p'})}\|u\|_p\|v\|_p\,, \tag{19.39}$$

which proves the assertion by density of C_c^∞ in L^p spaces. □

19.2. **Clever factorization of general kinetic energies.** Let h_0 be a tempered distribution on $\mathbb{R}^d$, smooth in a neighborhood of some point $\xi^0 \in \mathbb{R}^d$ such that $\lambda := h_0(\xi^0)$ is a regular value of h_0, i.e., that the level set

$$S_\lambda := \{\xi \in \mathbb{R}^d : h_0(\xi) = \lambda\} \tag{19.40}$$

is a smooth nonempty hypersurface near ξ^0. Locally, i.e., for ξ near ξ^0 the implicit function theorem implies that S_λ is the graph $\xi_1 = a(\xi', \lambda)$ for some real-valued function $a(\xi', \lambda)$. This yields the factorization

$$h_0(\xi) - \lambda = e(\xi, \lambda)(\xi_1 - a(\xi', \lambda)), \tag{19.41}$$

where $\xi = (\xi_1, \xi') \in \mathbb{R} \times \mathbb{R}^{d-1}$, and $a(\xi', \lambda)$, and $e(\xi, \lambda)$ are real-valued, smooth functions, $e(\xi, \lambda)$ being bounded away from zero. Locally, $e(\xi, \lambda)$ is given by the expression

$$e(\xi, \lambda) = \int_0^1 dt\, \partial_{\xi_1} h_0(t\xi_1 + (1-t)a(\xi', \lambda), \xi'). \tag{19.42}$$

By a linear change of coordinates one may always assume $a(0, \lambda) = 0$ and $\partial_{\xi'} a(\xi', \lambda)|_{\xi'=0} = 0$. Then $a(\xi', \lambda) = \mathcal{O}(|\xi'|^2)$.

Now suppose $I \subseteq \mathbb{R}$ is a fixed compact subset of the set of regular values of $h_0(\xi)$, i.e.,

$$\{\lambda \in \mathbb{R} : \nabla h_0(\xi) \neq 0 \text{ for all } \xi \in \mathbb{R}^d \text{ such that } h_0(\xi) = \lambda\}, \tag{19.43}$$

and let

$$S = \bigcup_{\lambda \in I} S_\lambda = \{\xi \in \mathbb{R}^d : h_0(\xi) \in I\}. \tag{19.44}$$

We assume that S_λ is compact and has everywhere non-vanishing curvature for each $\lambda \in I$. The following lemma is closely related to the Stein–Tomas theorem for the Fourier restriction operator.

Lemma 19.8 ([53, Lemma 3.3])**.** *Let η be a bump function. Then*

$$\sup_{\lambda \in I, |\varepsilon| \leq 1} \|\eta(D)[h_0(D) - (\lambda + i\varepsilon)]^{-1}\|_{p \to p'} \lesssim 1. \tag{19.45}$$

The following is a more precise version which tracks the precise ε-dependence.

Lemma 19.9 ([13, Lemma 23])**.** *Let η be a bump function, $\zeta \in \mathbb{C}$ such that $0 \leq \mathrm{Re}(\zeta) \leq (d+1)/2$, and $\varepsilon \in [-1, 1]$. Let also $R^{\eta,\zeta}_{\lambda,\varepsilon} = \eta(D)[h_0(D) - (\lambda + i\varepsilon)]^{-\zeta}$. Then we have for all $N \in \mathbb{N}$ the kernel bound*

$$\sup_{\lambda \in I} |R^{\eta,\zeta}_{\lambda,\varepsilon}(x)| \lesssim_N \mathrm{e}^{C|\mathrm{Im}(\zeta)|^2} \langle x \rangle^{-\frac{d-1}{2} + \mathrm{Re}(\zeta)} \langle \varepsilon x \rangle^{-N} \tag{19.46}$$

for some $C > 0$.

Proof. It suffices to prove this assertion for fixed λ. Thus, we absorb λ into the symbol and consider $p(\xi) = h_0(\xi) - \lambda$. Let $\Omega \subseteq \mathbb{R}^d$ be a precompact subset such that $S \Subset \Omega$. By a partition of unity and a linear change of coordinates we may assume that, locally near an arbitrary point of Ω, we have either $p \neq 0$, or $\partial p / \partial \xi_1 > 0$. In case $p \neq 0$ we get the stronger bound

$$|R^{\eta,\zeta}_{\lambda,\varepsilon}(x)| \lesssim_N \langle x \rangle^{-N}. \tag{19.47}$$

by non-stationary phase.

Suppose now $p = 0$ and $\partial p / \partial \xi_1 > 0$ and consider first $\zeta = 1$, i.e., we are dealing with the resolvent. By the implicit function theorem, the set $\{p(\xi) = 0\}$ is then the graph of a smooth function $\xi_1 = a(\xi')$ and we have the factorization

$$(p(\xi) - i\varepsilon)^{-1} = q(\xi)(\xi_1 - a(\xi') - i\varepsilon q(\xi))^{-1}, \tag{19.48}$$

where

$$q(\xi) = \frac{\xi_1 - a(\xi')}{p(\xi)} > 0\,, \tag{19.49}$$

see, e.g., [109, Section 14.2], [53, Lemma 3.3], or [175, Section 3.1]. There it is sufficient to work with the limiting distributions corresponding to $\varepsilon = \pm 0$, which would yield the assertion of the lemma in this case. However, we claim the estimates also for non-zero ε. To obtain the desired decay for fixed $\varepsilon > 0$, we use another factorization in which $q(\xi)$ will be replaced by a function which does not depend on ξ_1. The following approach is motivated by Koch and Tataru [124], albeit in the much simpler setting of constant coefficients. Their [124, Lemma 3.8] provides the alternative factorization

$$e(\xi)(p(\xi) - i\varepsilon) = \xi_1 + a(\xi') + i\varepsilon b(\xi')\,, \tag{19.50}$$

where e is elliptic ($e \neq 0$) and a, b are real-valued. This is a version of the Malgrange preparation theorem [111, Theorem 7.5.5] or the classical Weierstrass preparation theorem [111, Theorem 7.5.1] in the analytic case. We appeal to [124] because it makes the dependence on ε explicit. Note that the imaginary part b is now *independent of* ξ_1. The symbols e, a, b can be found by iteratively solving a system of algebraic equations and using Borel resummation of the resulting formal series (see [124, Lemmas 3.9 and 3.10]). Moreover, a, b have asymptotic expansions in powers of ε, while e has an asymptotic expansion in powers of ε and ξ_1. We will only need the first term b_1 in the expansion of b. Changing variables $\xi \mapsto \xi_1 + a(\xi')$ we are reduced to $p(\xi) = \xi + i\varepsilon q(\xi)$ for some real-valued function q. By the proof of [124, Lemma 3.9] we have $b_1 = 1/(1 + q_1^2)$, where $q_1 = \partial_{\xi_1} q|_{\xi_1=0}$. Therefore, $b \geq c$ on the closure of Ω for some constant $c > 0$ (we used compactness and the smallness of ε). Since we have constant coefficients, the simple parametrix [124, (5.5)] with (operator-valued) kernel $K(x_1 - y_1)$, given by

$$K(x_1) = \mathbf{1}_{\{x_1<0\}} \mathrm{e}^{\varepsilon x_1 b(D')} \mathrm{e}^{-ix_1 a(D')}\,, \tag{19.51}$$

is actually exact, i.e., we have $(D_1 + a(D') + i\varepsilon b(D'))^{-1} K$ is the identity (we denote both the operator and the kernel by K here). By the stationary phase estimate for complex-valued phase functions [111, Theorem 7.7.5], we have

$$|K(x)| \lesssim \langle x\rangle^{-\frac{d-1}{2}} \mathrm{e}^{-c|\varepsilon x|} + \mathcal{O}_N(\langle x\rangle^{-N})\,. \tag{19.52}$$

Using the factorization (19.50) and extending $1/e$ globally as a Schwartz function, we obtain the claimed estimate of the lemma in the case $\zeta = 1$.

The case $\zeta \neq 1$ requires only minor modifications. The kernel in (19.51) is replaced by

$$K_\zeta(x_1) = \chi_-^{\zeta-1}(x_1) \mathrm{e}^{\varepsilon x_1 b(D')} \mathrm{e}^{-ix_1 a(D')}\,, \tag{19.53}$$

where $\chi_-^w(\tau) := \mathbf{1}_{\{\tau<0\}} |\tau|^w / \Gamma(w+1)$, $w \in \mathbb{C}$, and Γ is the usual Gamma function. Then $(D_1 + a(D') + i\varepsilon b(D'))^{-\zeta} K_\zeta$ is the identity. This follows immediately by applying the inverse Fourier transform to the following identity (see, e.g., the explanation after [111, Example 7.1.17])

$$\mathcal{F}\left(\tau \mapsto \mathrm{e}^{-\delta\tau} \chi_+^\zeta(\tau)\right)(\xi) = \mathrm{e}^{-i\pi(\zeta+1)/2} (\xi - i\delta)^{-\zeta-1}\,, \quad \delta > 0\,, \quad \zeta \in \mathbb{C}\,. \tag{19.54}$$

Again by stationary phase,

$$|K_\zeta(x)| \lesssim \mathrm{e}^{C|\mathrm{Im}(\zeta)|^2} \left(\langle x\rangle^{-\frac{d-1}{2} + \mathrm{Re}(\zeta)} \mathrm{e}^{-c|\varepsilon x|} + \mathcal{O}_N(\langle x\rangle^{-N})\right)\,, \quad 0 \leq \mathrm{Re}(\zeta) \leq \frac{d+1}{2}\,. \tag{19.55}$$

The growth estimate in $|\mathrm{Im}(\zeta)|$ comes from a standard estimate on the Gamma function, cf. [96, Appendix A.7]. □

20. Stationary phase and microlocal analysis

We start with a classic review of the technique of stationary phase and apply it to obtain estimates on the Fourier transform of surface measures of curved, smooth surfaces. This material is classic and is covered exhaustively, e.g., in Stein [167, Chapter VIII]. Here, we will actually inspect the proofs a bit more closely and seek sufficient conditions on the smoothness of the manifold in question. Afterwards we will connect the stationary phase techniques to analyze certain distributions defined by oscillatory integrals and review the lattice counting problem. Then, we review some facts from pseudodifferential operators and microlocal analysis on $\mathbb{R}^d$ and transfer them to the setting of compact manifolds. Finally, we study the propagation of singularities and prove Egorov's theorem.

Concerning the first problem of obtaining bounds on $\hat{d\sigma}$, let $S \subset \mathbb{R}^d$ be a C^{N_φ} manifold of codimension one with non-vanishing Gaussian curvature and surface measure $d\sigma(\xi)$. Let $\psi \in C_c^{N_\psi}(\mathbb{R}^d)$ whose support intersects S in a compact subset of S. Denoting $d\mu = \psi d\sigma$, we wish to obtain the smallest $N_\varphi, N_\psi \in \mathbb{N}$ such that

$$|(d\mu)^\vee(x)| \lesssim < x >^{-(d-1)/2} \ .$$

In Proposition 20.5, we show that $N_\varphi \geq 4 + \lceil d/2 \rceil$ and $N_\psi \geq 2 + \lceil d/2 \rceil$ are sufficient conditions. Herz [105] showed that this regularity condition can even be relaxed to $N_\varphi \geq [(d-1)/2+2]$, if one sets $\psi = 1$ on S. (If $N_\varphi \geq [(d-1)/2+4]$, he obtained the leading term in the asymptotic expansion for $(d\sigma)^\vee$ as $|x| \to \infty$.)

The decay estimate for $(d\mu)^\vee$ is often proved using a stationary phase argument. Here, we follow the presentation of Stein [167, Chapter VIII] and start with a repetition on oscillatory integrals of the first kind.

20.1. **Oscillatory integrals of the first kind in one dimension.** In this section we consider integrals of the form

$$I(\lambda) := \int_{\mathbb{R}} \mathrm{e}^{i\lambda\varphi(x)} \psi(x)\, dx \tag{20.1}$$

for $\lambda \gg 1$, $\psi \in C_c^{N_\psi}(\mathbb{R})$, $\varphi \in C^{N_\varphi}(\mathbb{R})$, and certain $N_\varphi, N_\psi \in \mathbb{N}$.

Proposition 20.1. *Let $N \in \mathbb{N}$. If $\psi \in C_c^N(\mathbb{R})$ and $\varphi \in C^{N+1}(\mathbb{R})$ with $\varphi'(x) \neq 0$ on* suppψ*, then*

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

Proof. We define the "covariant derivative" D and its adjoint by

$$(Df)(x) := \frac{1}{i\lambda\varphi'(x)} f'(x) \quad \text{and} \quad ({}^tDf)(x) := -\frac{d}{dx}\left(\frac{f(x)}{i\lambda\varphi'(x)}\right) .$$

Since $D^N \mathrm{e}^{i\lambda\varphi} = \mathrm{e}^{i\lambda\varphi}$, integration by parts yields

$$\left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda\varphi(x)} \psi(x)\, dx\right| = \left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda\varphi(x)} ({}^tD)^N \psi(x)\, dx\right| \lesssim \lambda^{-N}$$

what was asserted. □

We will now consider the situation where φ' vanishes somewhere on suppψ. The case where also higher derivatives vanish can be found in [167, Chapter VIII, Proposition 3]. In particular, an asymptotic expansion is derived whose coefficients can be computed explicitly for certain phase functions φ, see also [167, Chapter VIII, Section 5.1].

Proposition 20.2. *Assume $\psi \in C_c^{N_\psi}(\mathbb{R})$, $\varphi \in C^{N_\varphi}(\mathbb{R})$ with $N_\psi \geq 3$ and $N_\varphi \geq 5$. Let $x_0 \in \mathrm{supp}\varphi$ be such that $\varphi(x_0) = \varphi'(x_0) = 0$, but $\varphi''(x_0) \neq 0$. Assume further that ψ is supported in a sufficiently small neighborhood around x_0. Then*

$$|I(\lambda)| \lesssim \lambda^{-1/2}\,.$$

Proof. We split the proof into four steps.

Step 1. We show that

$$\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \mathrm{e}^{-x^2}\,dx \sim \lambda^{-(\ell+1)/2} \sum_{j=0}^{\infty} c_j^{(\ell)} \lambda^{-j} \quad \ell \in \mathbb{N}_0\,. \tag{20.2}$$

The proof is contained in [167, Chapter VIII, Formula (9)].

Step 2. Let $\eta \in C_c^{\lceil(\ell+1)/2\rceil+1}(\mathbb{R})$. We will then show

$$\left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \eta(x)\,dx\right| \lesssim \lambda^{-(\ell+1)/2} \quad \ell \in \mathbb{N}_0\,. \tag{20.3}$$

To prove this, let $\alpha \in C^\infty$ with

$$\alpha(x) = \begin{cases} 1 & \text{for } |x| \leq 1\,, \\ 0 & \text{for } |x| \geq 2 \end{cases}$$

and decompose, for some $\varepsilon > 0$ to be chosen in a moment,

$$\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \eta(x)\,dx = \int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \eta(x)\alpha(x/\varepsilon)\,dx + \int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \eta(x)(1-\alpha(x/\varepsilon))\,dx$$

Clearly, the first summand is bounded by a constant times $\varepsilon^{\ell+1}$. To estimate the second summand, recall the covariant derivative D from Proposition 20.1 which, in this context, acts as

$$(Df)(x) = \frac{1}{2i\lambda x} f'(x) \quad \text{and} \quad ({}^tDf)(x) = \frac{i}{\lambda}\frac{d}{dx}\left(\frac{f(x)}{2x}\right)\,.$$

Thus, we have for $N > (\ell+1)/2$,

$$\begin{aligned} \left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} x^\ell \eta(x)(1-\alpha(x/\varepsilon))\,dx\right| &= \left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} ({}^tD)^N \left[x^\ell \eta(x)(1-\alpha(x/\varepsilon))\right]\,dx\right| \\ &\lesssim \lambda^{-N} \int\limits_{|x|\geq\varepsilon} |x|^{\ell-2N}\,dx = \mathrm{const}\; \lambda^{-N}\varepsilon^{\ell+1-2N}\,. \end{aligned}$$

Choosing $\varepsilon = \lambda^{-1/2}$ shows (20.3). Similarly, one obtains for any $g \in \mathcal{S}(\mathbb{R})$ vanishing near the origin,

$$\left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} g(x)\,dx\right| = \left|\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} ({}^tD)^N g(x)\,dx\right| \lesssim \lambda^{-N}\,, \quad N \in \mathbb{N}_0\,. \tag{20.4}$$

Step 3. We will now prove the assertion for $\varphi(x) = x^2$ and $\psi \in C_c^{N_\psi}(\mathbb{R})$ with $N_\psi \geq 3$. Let $\tilde{\psi} \in C_c^\infty(\mathbb{R}^d)$ with $\tilde{\psi}(x) = 1$ on $\mathrm{supp}\psi$, write

$$\int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} \psi(x)\,dx = \int_{\mathbb{R}} \mathrm{e}^{i\lambda x^2} \mathrm{e}^{-x^2} \left(\mathrm{e}^{x^2}\psi(x)\right) \tilde{\psi}(x)\,dx\,,$$

and Taylor expand to zeroth order

$$\mathrm{e}^{x^2}\psi(x) = b_0 + h_0(x) \cdot x$$

where $h_0(x) = o(x)$ belongs to $C^{N_\psi - 1}(\mathbb{R})$. Plugging this this into the above integral gives three terms, namely

$$b_0 \int_{\mathbb{R}} e^{i\lambda x^2} e^{-x^2}\, dx \sim b_0 \lambda^{-1/2} \sum_m c_m \lambda^{-m} \tag{20.5a}$$

$$\left| \int_{\mathbb{R}} e^{i\lambda x^2} x h_0(x) e^{-x^2} \tilde{\psi}(x)\, dx \right| \lesssim \lambda^{-1} \tag{20.5b}$$

$$\left| \int_{\mathbb{R}} e^{i\lambda x^2} b_0 e^{-x^2} \left(\tilde{\psi}(x) - 1 \right)\, dx \right| \lesssim \lambda^{-N} \tag{20.5c}$$

where we used (20.2) for the first summand, (20.3) for the second one (since $h_0 \in C^{N_\psi - 1}(\mathbb{R}) \subseteq C^2(\mathbb{R})$), and (20.4) for the third one.

Step 4. We finally consider general phase functions $\varphi \in C^{N_\varphi}(\mathbb{R})$ with $N_\varphi \geq 5$. We expand φ near x_0, i.e., $\varphi(x) = c(x-x_0)^2[1+\varepsilon(x)]$ for some $c \neq 0$ and $\varepsilon \in C^{N_\varphi - 2}(\mathbb{R})$ with $\varepsilon(x) = \mathcal{O}(|x-x_0|)$, i.e., $|\varepsilon(x)| \leq 1$ for x sufficiently close to x_0. For such x, one has in particular $\varphi'(x) \neq 0$. Thus, let us fix a neighborhood U around x_0 so small such that these conditions hold. Since we assumed that the support of ψ was small enough, we can in particular assume supp$\psi \subseteq U$. Now, let $y := (x - x_0)[1 + \varepsilon(x)]$, i.e., $x \mapsto y(x)$ is a $C^{N_\varphi - 2}(\mathbb{R})$ diffeomorphism from U to some neighborhood of the origin. Since $\varphi(x) = cy^2$, we have

$$\int_{\mathbb{R}} e^{i\lambda\varphi(x)} \psi(x)\, dx = \int_{\mathbb{R}} e^{ic\lambda y^2} \tilde{\psi}(y)\, dy$$

for some $\tilde{\psi} \in C^{N_\psi}(\mathbb{R}) \cap C^{N_\varphi - 2}(\mathbb{R})$ whose support intersects any neighborhood of the origin. Thus, we can apply the results of the third step and conclude the proof. □

20.2. **Oscillatory integrals of the first kind in higher dimensions.** We will now generalize Propositions 20.1 and 20.2 to $\mathbb{R}^d$ with $d \geq 2$. We will say that phase function φ defined in a neighborhood of a point $x_0 \in \mathbb{R}^d$ has x_0 as a *critical point* if

$$(\nabla\varphi)(x_0) = 0\,.$$

Similarly as before, let

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

Proposition 20.3. *Let $N \in \mathbb{N}$. If $\psi \in C_c^N(\mathbb{R}^d)$ and $\varphi \in C^{N+1}(\mathbb{R}^d)$ has no critical points in* suppψ*, then*

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

Proof. For each $x_0 \in$ suppψ there is a $\xi \in \mathbb{S}^{d-1}$ and a ball $B_{x_0}(\delta)$ for some $\delta \ll 1$ such that

$$\xi \cdot (\nabla\varphi)(x) \geq c > 0 \quad \text{for all } x \in B_{x_0}(\delta)\,.$$

Decompose $\psi = \sum_k \psi_k$ into a finite sum where each $\psi_k \in C_c^N(\mathbb{R}^d)$ is supported in one of these balls. Now choose a coordinate system $x_1, ..., x_d$ such that x_1 lies along ξ. Then

$$\int_{\mathbb{R}^d} e^{i\lambda\varphi(x)} \psi_k(x)\, dx = \int_{\mathbb{R}^{d-1}} dx_2 ... dx_d \left(\int_{\mathbb{R}} e^{i\lambda\varphi(x_1, ..., x_d)} \psi_k(x_1, ..., x_d)\, dx_1 \right)$$

and we can apply Proposition 20.1 to the x_1 integral to conclude the proof. □

Next, suppose φ has a critical point at x_0 but is *non-degenerate*. By that we mean that the $d \times d$ matrix

$$\frac{\partial^2 \varphi}{\partial x_j \partial x_k}$$

is invertible. Using a Taylor expansion (e.g., for $\varphi \in C^{N_\varphi}(\mathbb{R}^d)$ with $N_\varphi \geq 3$), one sees that non-degenerate critical points are in fact isolated.

Proposition 20.4. *Suppose $\varphi \in C^{N_\varphi}(\mathbb{R}^d)$ with $N_\varphi \geq 4 + \lceil (d+1)/2 \rceil$, and $x_0 \in \mathbb{R}^d$ is a non-degenerate, critical point of φ where additionally $\varphi(x_0) = 0$. If $\psi \in C_c^{N_\psi}(\mathbb{R}^d)$ with $N_\psi \geq 2 + \lceil (d+1)/2 \rceil$ is supported in a sufficiently small neighborhood of x_0, then*

$$|I(\lambda)| \lesssim \lambda^{-d/2}. \tag{20.6}$$

Moreover, for each $j = 1, 2, 3, ...$

$$\left| \partial_\lambda^j \left[\mathrm{e}^{-i\lambda\varphi(y_0)} I(\lambda) \right] \right| \lesssim_j \lambda^{-j-d/2}, \quad \lambda \geq 1 \tag{20.7}$$

and additionally

$$|I(\lambda)| \lesssim \lambda^{-1-d/2}, \quad \lambda \geq 1, \quad \text{if } \psi(y_0) = 0. \tag{20.8}$$

Proof. The proof follows closely the lines of that of Proposition 20.2. First, let $Q(x)$ denote the unit quadratic form given by

$$Q(x) = \sum_{j=1}^m x_j^2 - \sum_{j=m+1}^d x_j^2$$

for some fixed $m \in \{0, 1, ..., d\}$. The analogue of (20.2) is

$$\int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} \mathrm{e}^{-|x|^2} x^\ell \, dx \sim \lambda^{-d/2-|\ell|/2} \sum_{j=0}^\infty c_j(m,\ell) \lambda^{-j}, \tag{20.9}$$

whose proof can be found in [167, p. 345].

Next, the analogue of (20.3) is the statement that

$$\left| \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \, dx \right| \lesssim \lambda^{-d/2-|\ell|/2} \tag{20.10}$$

if $\eta \in C_c^{\lceil (\ell+d)/2 \rceil + 1}(\mathbb{R}^d)$. (As in the proof of Proposition 20.2, we will apply this estimate for $\ell = 1$ with $h_0(x)\tilde{\psi} \in C^{N_\psi - 1}(\mathbb{R}^d)$ in place of η, i.e., $N_\psi \geq 2 + \lceil (1+d)/2 \rceil$.) To prove it, we consider the cones

$$\Gamma_j := \{x \in \mathbb{R}^d : |x_j|^2 \geq |x'|^2/(2d)\}$$

and the smaller

$$\Gamma_j^0 := \{x \in \mathbb{R}^d : |x_j|^2 \geq |x'|^2/d\},$$

where $x' = (x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$. Then, since

$$\bigcup_{j=1}^d \Gamma_j^0 = \mathbb{R}^d,$$

we can find functions $\Omega_1, ..., \Omega_d$ with $\mathrm{supp}\Omega_j \subseteq \Gamma_j$ which are homogeneous of degree zero and smooth away from the origin such that

$$\sum_{j=1}^d \Omega_j(x) = 1 \quad \text{for all } x \neq 0.$$

Thus, we can write

$$\int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \, dx = \sum_{j=1}^d \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \Omega_j(x) \, dx.$$

Now, as in the proof of (20.3), let $\alpha \in C_c^\infty(\mathbb{R}^d)$ be a radial function such that

$$\alpha(x) = \begin{cases} 1 & \text{for } |x| \le 1\,, \\ 0 & \text{for } |x| \ge 2\,, \end{cases}$$

and decompose

$$\begin{aligned} \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \Omega_j(x)\, dx = &\int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \Omega_j(x) \alpha(x/\varepsilon)\, dx \\ &+ \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \Omega_j(x) (1-\alpha(x/\varepsilon))\, dx\,. \end{aligned}$$

As before, the first summand is bounded by a constant times $\varepsilon^{\ell+d}$. To treat the second summand, we integrate by parts in the cone Γ_j, using the covariant derivative

$$D_j \mathrm{e}^{i\lambda Q(x)} = \mathrm{e}^{i\lambda Q(x)} \quad \text{with} \quad (D_j f)(x) = \pm \frac{1}{2i\lambda x_j} \frac{\partial f(x)}{\partial x_j}\,.$$

This, together with the fact that $|x_j| \ge |x'|/\sqrt{2d}$ in Γ_j, and

$$|({}^t D_j)^N \Omega_j(x)| \lesssim_N \lambda^{-N} |x|^{-2N}\,,$$

allows us to estimate

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} x^\ell \eta(x) \Omega_j(x) (1-\alpha(x/\varepsilon))\, dx \right| \\ &\quad = \left| \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} ({}^t D_j)^N \left[x^\ell \eta(x) \Omega_j(x) (1-\alpha(x/\varepsilon)) \right]\, dx \right| \\ &\lesssim \lambda^{-N} \int_{|x|\ge\varepsilon, |x_d| \ge |x'|/\sqrt{2d}} |x|^{\ell-2N}\, dx \lesssim \lambda^{-N} \varepsilon^{\ell-2N+d} \end{aligned}$$

for $N > (\ell+d)/2$. Choosing $\varepsilon = \lambda^{-1/2}$ as before shows (20.10).

A similar argument shows that whenever $g \in \mathcal{S}(\mathbb{R}^d)$ and g vanishes near the origin, then

$$\left| \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda Q(x)} g(x)\, dx \right| \lesssim \lambda^{-N}\,, \quad N \in \mathbb{N}_0\,, \tag{20.11}$$

which is the analog of (20.4). Combining this with (20.9) and (20.10) as in the proof of Proposition 20.2 yields the assertion in the special case $\varphi(x) = Q(x)$.

To pass to the general case, one can appeal to the change of variables guaranteed by Morse's lemma. Since $\varphi(x_0) = \nabla\varphi(x_0) = 0$, and the critical point is assumed to be non-degenerate, there exists a $C^{N_\varphi - 2}(\mathbb{R}^d)$ diffeomorphism from a small neighborhood of x_0 in x-space to a small neighborhood of the origin in y-space under which φ is transformed into

$$\sum_{j=1}^m y_j^2 - \sum_{j=m+1}^d y_j^2\,,$$

for some $m \in \{0, ..., d\}$. The index m is the same as that of the quadratic form corresponding to

$$\left[\frac{\partial^2 \varphi}{\partial x_j \partial x_k} \right] (x_0)\,.$$

The proof of this can found in [167, p. 346-347]. Combining this with the findings in the special case where $\varphi(x) = Q(x)$, concludes the proof. ☐

20.3. **Fourier transforms of measures supported on surfaces.** Let $\varphi \in C_c^{N_\varphi}(\mathbb{R}^n)$ with $N_\varphi \geq 4 + \lceil (n+1)/2 \rceil$, and $\varphi(0) = \nabla\varphi(0) = 0$. Let us further assume that the determinant of the $n \times n$ matrix

$$\left(\frac{\partial^2 \varphi}{\partial \xi_j \partial \xi_k}\right)(\xi = 0)$$

never vanishes. Then φ describes a n-dimensional C^{N_φ} surface S, which is given by the graph $\xi_{n+1} = \varphi(\xi_1, ..., \xi_n)$ and has non-zero Gaussian curvature at every point. Let $d\sigma$ denote the measure on S induced by the Lebesgue measure on $\mathbb{R}^{n+1}$, and fix a function $\psi \in C_c^{N_\psi}(\mathbb{R}^{n+1})$ with $N_\psi \geq 2 + \lceil (n+1)/2 \rceil$ whose support intersects S in a compact subset of S. Let us now consider the finite Borel measure $d\mu(\xi) = \psi(\xi) d\sigma(\xi)$ on $\mathbb{R}^{n+1}$, which is of course carried on S. We wish to discuss the behavior of the Fourier transform

$$(d\mu)^\vee(x) = \int_S \mathrm{e}^{2\pi i x \cdot \xi} \psi(\xi)\, d\sigma(\xi)$$

for large $|x|$. For convenience, we relabel $d = n + 1$ in the following

Proposition 20.5. *Suppose S is a C^{N_φ} surface in $\mathbb{R}^d$ of codimension one with $N_\varphi \geq 4 + \lceil d/2 \rceil$, whose Gaussian curvature is non-zero everywhere. Let further $d\mu = \psi d\sigma$ be as above. Then*

$$|(d\mu)^\vee(x)| \lesssim |x|^{-(d-1)/2}\,.$$

Proof. For the purpose of the proof (in applying Proposition 20.4), we will work with $n = d - 1$ as in the beginning of this section and assume, by compactness, that S is given by the graph

$$\xi_{n+1} = \varphi(\xi_1, ..., \xi_d)\,,$$

so $d\sigma(\xi) = \sqrt{1 + |\nabla\varphi(\xi)|^2} d\xi_1 ... d\xi_n$. Thus, we can reduce matters to showing that, if $\tilde{\psi} \in C_c^{N_\psi}(\mathbb{R}^n)$ with $N_\psi \geq 2 + \lceil (n+1)/2 \rceil$ is supported in a small neighborhood of the origin,

$$\left| \int_{\mathbb{R}^n} \mathrm{e}^{i\lambda\Phi(\xi,\eta)} \tilde{\psi}(\xi)\, d\xi \right| \lesssim \lambda^{-n/2} \tag{20.12}$$

where $\lambda = |x| > 0$, $x = \lambda\eta$, and $\eta = (\eta_1, ..., \eta_{n+1})$ is a unit vector, and

$$\Phi(\xi, \eta) = \xi \cdot \eta = \sum_{j=1}^n \xi_j \eta_j + \varphi(\xi_1, ..., \xi_n)\eta_{n+1}\,.$$

Also, we have that $\varphi(0) = \nabla\varphi(0) = 0$, and

$$\det_{1 \leq j,k \leq n} \left(\frac{\partial^2 \varphi}{\partial \xi_j \partial \xi_k}\right)(0) \neq 0\,.$$

We divide the proof into three cases, depending on the position of $\eta \in \mathbb{S}^n$, namely

(1) η is sufficiently close to the "north pole" $\eta_N = (0, 0, ..., 1)$,
(2) η is sufficiently close to the "south pole" $\eta_S = (0, 0, ..., -1)$, and
(3) η lies in the complementary set on the unit sphere.

The first and second case are analogous. We have that $\nabla_\xi \Phi(\xi, \eta_N)|_{\xi=0} = 0$ and want to see that for each η sufficiently close to η_N, there is a unique $\xi = \xi(\eta)$ so that

$$\nabla_\xi \Phi(\xi, \eta)|_{\xi=\xi(\eta)} = 0\,.$$

The latter is a series of n equations, and one can find the desired solution by the implicit function theorem, which requires that we check that the Jacobian determinant

$$\det \left[\frac{\partial^2 \varphi}{\partial \xi_j \partial \xi_k}\right](0, \eta_N) \neq 0\,,$$

but this is of course our assumption of the non-vanishing curvature. In particular, if the η-neighborhood of η_N is sufficiently small, then also

$$\det\left[\frac{\partial^2\varphi}{\partial\xi_j\partial\xi_k}\right](\xi(\eta),\eta)\neq 0\,,$$

and we can invoke Proposition 20.4 (with $x_0=\xi(\eta)$) as long as the support of $\tilde{\psi}$ is small enough. This shows that the left side of (A.8) is bounded by a constant times $\lambda^{-n/2}$ and concludes the discussion in the first two cases.

Thus, we are left with the third class of η. By definition,

$$\nabla_\xi\Phi(\xi,\eta)=(\eta_1,...,\eta_n)+\eta_{n+1}\nabla\varphi(\xi)\,.$$

However, $(\eta_1^2+...+\eta_n^2)^{1/2}\geq c>0$ for η away from the poles, and

$$\nabla\varphi(\xi)=\mathcal{O}(\xi)\quad\text{as }\xi\to 0\,.$$

Thus, $|\nabla_\xi\Phi(\xi,\eta)|\geq c'>0$, if the support of $\tilde{\psi}$ is a sufficiently small neighborhood of the origin. We may now invoke Proposition 20.3 (with $N=2+\lceil(n+1)/2\rceil$) which shows that the left side of (A.8) is bounded by a constant times $\lambda^{-2-\lceil(n+1)/2\rceil}\leq\lambda^{-n/2}$. □

20.4. **Oscillatory integrals and wave front sets.** Here, we follow Sogge [160, Section 4.1.1] but refer also to the classic exposition of Hörmander [111, Section 7.8 and Chapter VIII].

We now apply the "nonstationary phase lemma" (Proposition 20.4) to analyze certain distributions defined by oscillatory integrals. Specifically, let us consider integrals of the form

$$I_\Phi(x)=\int_{\mathbb{R}^N}\mathrm{e}^{i\Phi(x,\theta)}a(x,\theta)\,d\theta\equiv\lim_{\varepsilon\to 0}\int_{\mathbb{R}^N}\mathrm{e}^{i\Phi(x,\theta)}a(x,\theta)\rho(\varepsilon\theta)\,d\theta \tag{20.13}$$

where in this definition $\rho\in C_c^\infty(\mathbb{R}^N)$ is a bump that equals one near the origin. In fact, for the oscillatory integrals that we consider here, we will see that the definition does not depend on the particular choice of ρ.

Here, we assume $x\in\Omega\subseteq\mathbb{R}^d$ where Ω is an open subset of $\mathbb{R}^d$ with d possibly different from N. Moreover, we assume $\Phi\in C^\infty(\Omega\times\mathbb{R}^N\setminus\{0\})$ is real, homogeneous of degree one, i.e.,

$$\Phi(x,\lambda\theta)=\lambda\Phi(x,\theta)\,,\quad\lambda>0 \tag{20.14}$$

and, additionally, if d denotes the differential with respect to all variables, we assume

$$d\Phi\neq 0\quad\text{on }\Omega\times\mathbb{R}^N\setminus\{0\}\,. \tag{20.15}$$

As an example, one may think of $\Phi(x,\theta)=x'\cdot\theta+x_{N+1}\theta^2$ with $x=(x',x_{N+1})\in\mathbb{R}^{N+1}$ and $\xi\in\mathbb{R}^N$. Finally, we shall also assume that the amplitude $a(x,\theta)$ is a standard symbol of order m, i.e., for all multi-indices α and γ, we have

$$|D_x^\gamma D_\theta^\alpha a(x,\theta)|\lesssim_{\alpha,\gamma}(1+|\theta|)^{m-|\alpha|}\,, \tag{20.16}$$

whenever x belongs to a fixed compact subset of Ω and $\theta\in\mathbb{R}^N$. In this case, we shall abbreviate

$$a\in S^m\Leftrightarrow\text{(20.16) is valid.}$$

We will now give a sufficient condition when I_Φ in (20.13) is smooth.

Theorem 20.6. *If Φ is as above and $a\in S^m$, then $I_\Phi\in\mathcal{D}'(\Omega)$ and its definition* (20.13) *does not depend on the choice of ρ. Additionally, if $x_0\in\Omega$ and*

$$\nabla_\theta\Phi(x_0,\theta)\neq 0\ \textit{for all}\ \theta\in\mathbb{R}^N\setminus\{0\}\,,$$

then I_Φ is smooth in a neighborhood of x_0.

Before we turn th the proof, we restate the last part of the theorem. We recall

Definition 20.7. Let $v \in \mathcal{D}'(\Omega)$. Then the *singular support* $\operatorname{sing\,supp} v$ of v is defined as the complement of the set of points $x_0 \in \Omega$ which have the property that v restricts as an element of $C^\infty(\mathcal{N}_{x_0})$ for some neighborhood $\mathcal{N}_{x_0}$ of x_0.

Using this notion, the last part of Theorem 20.6 says

$$\operatorname{sing\,supp} I_\Phi \subseteq \{x \in \Omega : \nabla_\theta \Phi(x,\theta) = 0 \text{ for some } \theta \in \mathbb{R}^N \setminus \{0\}\} . \tag{20.17}$$

Proof. We first show $I_\Phi \in \mathcal{D}'(\Omega)$. To do so, we decompose I_Φ dyadically. So, let $\beta \in C_c^\infty(\mathbb{R}^N)$ be a bump function with

$$\beta(\theta) = 0 \text{ if } |\theta| \notin [1/2, 2] , \quad \text{and} \quad \sum_{j=-\infty}^{\infty} \beta(\theta/2^j) = 1 , \quad \theta \neq 0 .$$

We then define for $u \in C_c^\infty(\Omega)$,

$$I_\Phi^j[u] = \int_\Omega dx \int_{\mathbb{R}^N} d\theta\ \mathrm{e}^{i\Phi(x,\theta)} \beta(\theta/2^j) a(x,\theta) u(x)$$

and

$$I_\Phi^0[u] = \int dx \int d\theta\ \mathrm{e}^{i\Phi(x,\theta)} \left(1 - \sum_{j=1}^{\infty} \beta(\theta/2^j)\right) a(x,\theta) u(x) .$$

Clearly, each I_Φ^j is a distribution on Ω for $j = 0, 1, 2, ...$ (which is just integration against a smooth function on Ω, depending on j). To prove that also I_Φ belongs to $\mathcal{D}'(\Omega)$, we show, for a given relatively compact subset $K \Subset \Omega$ and a number $M \in \mathbb{N}$ there is $k(M)$ such that

$$|I_\Phi^j[u]| \lesssim_M 2^{-Mj} \sup_{|\alpha| \le k(M)} \sup |D^\alpha u| , \quad u \in C^\infty(K) \quad \text{for all } j = 1, 2, ... \tag{20.18}$$

Setting $\lambda = 2^j$, one obtains

$$I_\Phi^j[u] = \lambda^N \iint \mathrm{e}^{i\Phi(x,\theta)} \beta(\theta/2^j) a(x,\theta) u(x)\, d\theta\, dx .$$

But since $a \in S^m$, we have

$$|D_x^\gamma D_\theta^\alpha (\beta(\theta) a(x, \lambda\theta))| \lesssim_{\alpha,\gamma,K} \lambda^m , \quad \text{for } x \in K .$$

Consequently, (20.18) follows from stationary phase (Proposition 20.3) and the assumption $d\Phi \neq 0$. Moreover, (20.18) implies that the definition (20.13) is indeed independent of ρ since we assumed that $\rho \in C_c^\infty(\mathbb{R}^N)$ equals one near the origin; consequently, if $\tilde\rho$ were another function with this property, then $\tilde\rho - \rho \in C_c^\infty(\mathbb{R}^N \setminus \{0\})$.

To prove (20.17), let $x_0 \in \Omega$ have the property that $\nabla_\theta \Phi(x_0, \theta) \neq 0$ for $\theta \in \mathbb{R}^N \setminus \{0\}$. We will now show that there is a $\delta > 0$ such that $I_\Phi(x)$ is smooth on $\{x : |x - x_0| < \delta\}$. Since Φ is homogeneous of degree one, we see that there is $\delta > 0$ and $c > 0$ such that

$$|\nabla_\theta \Phi(x,\theta)| \ge c \quad \text{if } |x - x_0| < \delta .$$

Therefore, if

$$L := \frac{\nabla_\theta \Phi(x,\theta)}{i\lambda |\nabla_\theta \Phi(x,\theta)|} \cdot \nabla_\theta ,$$

we have that for every M and $\{x : |x - x_0| \le \delta\}$ that

$$I_\Phi^j(x) = \lambda^N \int_{\mathbb{R}^N} \mathrm{e}^{i\lambda\Phi(x,\theta)} (L^*)^M (\beta(\theta) a(x, \lambda\theta))\, d\theta = \mathcal{O}(\lambda^{N+m-M}) .$$

Thus, if $M > N + m$ and $\chi(\theta) = \sum_{j=1}^{\infty} \beta(\theta/2^j)$, then

$$I_\Phi(x) - I_\Phi^0(x) = \int_{\mathbb{R}^N} e^{i\Phi(x,\theta)} (L^*)^M (\chi(\theta) a(x,\theta))\, d\theta$$

is an absolutely convergent integral. But that shows that $I_\Phi - I_\Phi^0$ is continuous on $\{x : |x - x_0| \leq \delta\}$ and, by similar arguments, that this difference is indeed smooth on this set. Since $I_\Phi^0 \in C^\infty(\Omega)$, this shows (20.17). □

While this theorem locates the possible locations of the singularities of I_Φ, it does not yet assert anything about the "directions of propagation" of these singularities.

Example 20.8. Let $x' = (x_1, ..., x_{d-1})$ and $\delta_0(x_d) = dx'$ be the induced Lebesgue measure on the hyperplane $x_d = 0$. Then the distributions $v = \rho dx'$ with $\rho \in C_c^\infty(\mathbb{R}^d)$ satisfy $\operatorname{sing\,supp} v \subseteq \operatorname{supp}\rho \cap \{x \in \mathbb{R}^d : x_d = 0\}$. On the other hand, since $\delta_0(x_d)$ is a distribution that does not depend on the x' variables, the "directions of the singularities of v" is just those spanned by the unit vectors $(0, ..., 0, \pm 1)$. (We will make this saying precise below.) This fact is captures by the Fourier transform, $\hat{v}(\xi)$, which is rapidly decreasing in any closed cone through the origin which does not contain $(0, ..., 0, \pm 1)$. For a generalization of this example, see Hörmander [111, Theorem 8.1.5].

Let us now consider more general $u \in \mathcal{E}'(\mathbb{R}^d)$ (compactly supported distributions). By a Paley–Wiener–Schwartz theorem, we have $u \in C_c^\infty(\mathbb{R}^d)$ if and only if $\hat{u}(\xi)$ is rapidly decreasing, i.e., $|\hat{u}(\xi)| \lesssim_N < \xi >^{-N} e^{H(\operatorname{Im}(\xi))}$ for any $\xi \in \mathbb{C}^d$ where $H(\xi) = \sup_{x \in \operatorname{supp} u} \langle x, \xi \rangle$ is the supporting function (see Sogge [160, §A.2] or Hörmander [111, Theorem 7.3.1]). However, the above example indicates that it is possible that $\hat{u} \in C^\infty$ is rapidly decreasing in some directions but not in the others, i.e., only *some* high-frequency components of $\hat{u}$ may contribute to the singularities of u. The *wave front set*, which we are about to define unifies these along with the singular support. Recall that a conic neighborhood of a set $\Sigma \subseteq \mathbb{R}^d \setminus \{0\}$ is an open set $\mathcal{N}$ containing Σ and having the property that if $\xi \in \mathcal{N}$, then so is $\lambda\xi$ for every $\lambda > 0$.

Definition 20.9. For $u \in \mathcal{E}'$ let $\Gamma(u) \subseteq \mathbb{R}^d \setminus \{0\}$ be the closed cone consisting of all $\eta \in \mathbb{R}^d \setminus \{0\}$ such that η has no conic neighborhood in which

$$|\hat{u}(\xi)| \lesssim_N < \xi >^{-N}, \quad N \in \mathbb{N}$$

holds.

Note that if $u \in \mathcal{E}'(\mathbb{R}^d)$, then, by Paley–Wiener, we have $u \in C_c^\infty$ if and only if $\Gamma(u) = \emptyset$. We may therefore interpret $\operatorname{sing\,supp} u$ as measuring the *location of the singularities of* u and $\Gamma(u)$ as measuring the *the directions of the singularities of* u. Keeping this in mind, we have the following natural result.

Lemma 20.10. *If $\rho \in C_c^\infty(\mathbb{R}^d)$ and $u \in \mathcal{E}'(\mathbb{R}^d)$, then*

$$\Gamma(\rho u) \subseteq \Gamma(u)\,.$$

Proof. Our goal is to control

$$\widehat{\rho u}(\xi) = \int \hat{\rho}(\xi - \eta)\hat{u}(\eta)\, d\eta\,.$$

Since $u \in \mathcal{E}'(\mathbb{R}^d)$, we know that $\hat{u}$ is smooth and satisfies

$$|\hat{u}(\eta)| \lesssim (1 + |\eta|)^m$$

for some m (by integration by parts, see also [111, Theorem 7.3.1]). Next, we note that if ξ is outside of a fixed conic neighborhood of $\Gamma(u)$ and η is inside a slightly smaller conic neighborhood, then $|\xi-\eta|\geq c(|\xi|+|\eta|)$ for some $c>0$. In this case, we obtain

$$|\hat{\rho}(\xi-\eta)\hat{u}(\eta)|\lesssim_N (1+|\xi|+|\eta|)^{-N}(1+|\eta|)^m \lesssim_N (1+|\xi|+|\eta|)^{-N+m}\,, \quad N\in\mathbb{N}\,.$$

On the other hand, if η is outside of a fixed small conic neighborhood of $\Gamma(u)$, we obtain for any $\xi\in\mathbb{R}^d$,

$$|\hat{\rho}(\xi-\eta)\hat{u}(\eta)|\lesssim_N (1+|\xi-\eta|)^{-N}(1+|\eta|)^{-N}\,.$$

Combining these two observations gives

$$\begin{aligned}|\widehat{\rho u}| &\lesssim_N \int (1+|\xi|+|\eta|)^{-N+m}\,d\eta + \int (1+|\xi-\eta|)^{-N}(1+|\eta|)^{-N}\,d\eta \\ &= \mathcal{O}(|\xi|^{-N+m+d}+|\xi|^{-N+d})\,,\end{aligned}$$

thereby showing $\Gamma(\rho u)\subseteq\Gamma(u)$. □

This lemma affords us a further localization.

Definition 20.11. Let $\Omega\subseteq\mathbb{R}^d$ be open and $u\in\mathcal{D}'(\Omega)$. For $x\in\Omega$, let

$$\Gamma_x(u) := \bigcap_{\{\rho\in C_c^\infty:\,\rho(x)\neq 0\}} \Gamma(\rho u)\,.$$

One easily verifies $\Gamma(\rho_j u)\to\Gamma_x(u)$ if ρ_j is a sequence of $C_c^\infty(\Omega)$ functions with $\rho_j(x)\neq 0$ and $\operatorname{supp}\rho_j\to\{x\}$, see also [111, pp. 253-254]. The set $\Gamma_x(u)\subseteq\mathbb{R}^d\setminus\{0\}$ essentially captures the directions of the singularities of u at x. This allows us to define a basic object in microlocal analysis.

Definition 20.12 (Wave front set)**.** For $u\in\mathcal{D}'(\Omega)$, the wave front set of u is defined as

$$WF(u) := \{(x,\xi)\in\Omega\times\mathbb{R}^d\setminus\{0\}:\,\xi\in\Gamma_x(u)\}\,.$$

Since $u\in\mathcal{D}'(\Omega)$ is smooth near x if and only if $\Gamma_x(u)=\emptyset$ (by Paley–Wiener), it follows that the projection of $WF(u)$ onto Ω is exactly $\operatorname{sing\,supp} u$. Similarly, one shows (see also [111, Proposition 8.1.2]) that the projection of $WF(u)$ onto the frequency component is precisely $\Gamma(u)$. In particular, this shows that $WF(u)$ is conic in the sense that it is invariant under multiplication by positive scalars in the second variable. It could therefore be considered as a subset of $\Omega\times\mathbb{S}^{d-1}$.

Theorem 20.13. *Let Ω be a linear subspace of $\mathbb{R}^d$ and $u=u_0\,d\Sigma$ where $u_0\in C^\infty(\Omega)$ and $d\Sigma$ is the Euclidean surface measure. Then*

$$WF(u)=\operatorname{supp} u\times(\Omega^\perp\setminus\{0\})\,.$$

As an example, think of $u=u_0\,dx'$, i.e., where $dx'=\delta(x_d)\,dx$ and $\Omega=\{x\in\mathbb{R}^d:\,x_d=0\}$.

Proof. See Hörmander [111, Theorem 8.1.5]. □

The following theorem naturally extends Theorem 20.6 and gives a first localization of $WF(I_\Phi)$.

Theorem 20.14. *Let $I_\Phi\in\mathcal{D}'(\Omega)$ be as in (20.13). Then*

$$WF(I_\Phi)\subseteq\{(x,\nabla_x\Phi(x,\theta)):\,(x,\theta)\in\Omega\times\mathbb{R}^N\setminus\{0\}\text{ and }\nabla_\theta\Phi(x,\theta)=0\}\,. \tag{20.19}$$

Proof. The proof is very similar to the one of Theorem 20.6. Let $u\in C_c^\infty(\Omega)$. To prove (20.19), it therefore suffices to show that

$$I(\xi) := \iint \mathrm{e}^{i\Phi(x,\theta)-ix\cdot\xi}u(x)a(x,\theta)\,d\theta\,dx$$

is rapidly decreasing when ξ is outside of an open cone Γ_0 containing

$$\{\nabla_x\Phi(x,\theta):\,(x,\theta)\in\operatorname{supp} u\times\mathbb{R}^N\setminus\{0\},\,\nabla_\theta\Phi(x,\theta)=0\}\,.$$

Repeating the previous arguments, this amounts to showing that for such ξ we have

$$\left|\iint e^{i\lambda\Phi(x,\theta)-ix\cdot\xi}u(x)\beta(\theta)a(x,\lambda\theta)\,dx\,d\theta\right| \lesssim_M (\lambda+|\xi|)^{-M}\,,\quad M\in\mathbb{N} \tag{20.20}$$

whenever $\beta\in C_c^\infty(\mathbb{R}^N\setminus\{0\})$. Let us define

$$\Psi(x,\theta) := \frac{\lambda\Phi(x,\theta)-x\cdot\xi}{\lambda+|\xi|}\,.$$

Then we claim that

$$|\nabla_{x,\theta}\Psi(x,\theta)| \sim \frac{|\lambda\nabla_x\Phi(x,\theta)-\xi|+\lambda|\nabla_\theta\Phi(x,\theta)|}{\lambda+|\xi|} \geq c>0\,,\quad \xi\notin\Gamma_0\,, \tag{20.21}$$

on the support of $u(x)\beta(\theta)a(x,\lambda\theta)$. This would show (20.20) after an application of the nonstationary phase lemma (Proposition 20.3).

To verify that claim, first note that (20.21) clearly holds, unless

$$c\leq\lambda/|\xi|\leq C$$

for certain constants $0<c<C<\infty$, since $d\Phi\neq 0$. So let us assume this in the following. Also, if $\nabla_\theta\Phi=0$, then $|\lambda\nabla_x\Phi(x,\theta)-\xi|\geq c'(|\lambda\nabla_x\Phi(x,\theta)|+|\xi|)$ for some $c'>0$ if ξ is outside of Γ_0. Thus, the claim holds when $|\nabla_\theta\Phi(x,\theta)|$ is small and $\beta(\theta)\neq 0$. Since (20.21) also clearly holds for such θ when $|\nabla\Phi(x,\theta)|$ is bounded from below, the proof is complete. □

We conclude this subsection by showing that $WF(u)$ is invariant under diffeomorphisms. Let

$$\kappa:\,\Omega\to\tilde{\Omega}$$

be a diffeomorphism between two open sets. Then if, say, u is a $L^1_{loc}(\tilde{\Omega})$ function, then it defines a distribution in $\mathcal{D}'(\tilde{\Omega})$, defined via

$$u(\Psi) := \int_{\tilde{\Omega}} u(y)\Psi(y)\,dy\,,\quad \Psi\in C_c^\infty(\tilde{\Omega})\,.$$

Likewise, the pullback of u via κ, i.e., $(\kappa^*u)(x)\equiv u(\kappa(x))$, defines an element of $\mathcal{D}'(\Omega)$. In this case, if $\psi\in C_c^\infty(\Omega)$, we get

$$(\kappa^*u)(\psi) = \int_\Omega u(\kappa(x))\psi(x)\,dx = \int_{\tilde{\Omega}} u(y)\psi(\kappa^{-1}(y))\left|\det\frac{d\kappa^{-1}}{dy}(y)\right|\,dy\,,\quad \psi\in C_c^\infty(\Omega)\,.$$

To be consistent, we must then define the pullback of a general $u\in\mathcal{D}'(\tilde{\Omega})$ by the formula

$$(\kappa^*u)(\psi) = u(\Psi)\,,\quad \Psi(y) = \psi(\kappa^{-1}(y))\left|\det\frac{d\kappa^{-1}}{dy}(y)\right|\,dy\,,\quad \psi\in C_c^\infty(\Omega)\,. \tag{20.22}$$

Note that if $\kappa:\,\mathbb{R}^d\to\mathbb{R}^d$ is a linear transformation, then (20.22) immediately gives the change of variables formula for wave front sets, i.e.,

$$WF(\kappa^*u) = \kappa^*WF(u)\,,\quad u\in\mathcal{D}'(\tilde{\Omega})\,, \tag{20.23}$$

whenever the pullback of a subset $\Lambda\subseteq\tilde{\Omega}\times\mathbb{R}^d\setminus\{0\}$ is defined via the pullback map for cotangent bundles, i.e.,

$$\kappa^*\Lambda := \{(x,\xi)\,:\,(\kappa(x),({}^t\kappa')^{-1}\xi)\in\Lambda\}\,. \tag{20.24}$$

The following result says that this fact remains true for general diffeomorphisms.

Theorem 20.15. *Let $\kappa:\Omega\to\tilde{\Omega}$ be a diffeomorphism between two open subsets of $\mathbb{R}^d$. Then* (20.23) *is valid.*

Remark 20.16. Note that the pullback formula (20.24) is exactly the change of variables for the cotangent bundle that one encounters in dealing with C^∞ manifolds. Thus, if M is a smooth d-dimensional manifold and $u \in \mathcal{D}'(M)$, then its wave front set $WF(u)$ can be defined as a subset of $T^*M \setminus \{0\}$ using local coordinates.

Proof of Theorem 20.15. □

20.5. **The lattice counting problem.** The goal of this section is to prove a primitive result concerning lattice counting in $\mathbb{R}^d$. Specifically, we show that

$$\#\{j \in \mathbb{Z}^d : |j| \le \lambda\} = |B_0(1)|\lambda^d + \mathcal{O}(\lambda^{d-2+\frac{2}{d+1}}), \quad \lambda \ge 1. \tag{20.25}$$

Using the decay of the Fourier transform of surface measures (Proposition 20.5) which in particular applies to the sphere, we obtain the following estimate on the Fourier transform of the ball multiplier.

Corollary 20.17. *Let $\chi(x)$ denote the characteristic function of the unit ball in $\mathbb{R}^d$, i.e., $\chi(x) = \mathbf{1}_{B_0(1)}(x)$. Then it satisfies*

$$|\hat{\chi}(\xi)| \lesssim < \xi >^{-\frac{d+1}{2}} . \tag{20.26}$$

Proof. First, since $\chi(x)$ is compactly supported, its Fourier transform is bounded (in fact even real analytic), i.e., it suffices to consider $|\xi| \ge 1$, say. Next, we reduce the problem to that region where $\chi(x)$ lacks continuity, i.e., an annulus around the unit sphere. For that purpose, let

$$C^\infty(\mathbb{R}) \ni \beta(r) := \begin{cases} 0 & \text{for } r \le 1/4 \\ 1 & \text{for } r \ge 1 \end{cases}$$

and smooth in $[1/4, 1]$. Then $(1 - \beta(|x|))\chi(x) \in C_c^\infty(\mathbb{R}^d)$, i.e., it has rapidly decaying Fourier transform. Thus, it suffices to prove

$$\int_{\mathbb{R}^d} \chi(x)\beta(|x|)\mathrm{e}^{-2\pi i x\cdot\xi}\, dx = \int_{1/4}^1 dr\ r^{d-1}\beta(r) \int_{\mathbb{S}^{d-1}} \mathrm{e}^{-2\pi i r\omega\cdot\xi} d\sigma(\omega) = \mathcal{O}(|\xi|^{-(d+1)/2})$$

where $d\sigma(\omega)$ denotes the usual Lebesgue measure on $\mathbb{S}^{d-1}$. We already saw that the Fourier transform of measures supported on curved surfaces is of the form (see, e.g., Stein [167, p. 360] or Sogge [160, Theorem 4.1.10])

$$\sum_{\pm} \mathrm{e}^{\pm 2\pi i r|\xi|} a_\pm(r|\xi|)\,,$$

where

$$\frac{d^j}{ds^j} a_\pm(s) = \mathcal{O}(s^{-\frac{d-1}{2}-j})\,, \quad j = 0, 1, 2, \ldots, \quad s > 1\,.$$

Plugging this in and integrating by part gives

$$\int_{\mathbb{R}^d} \chi(x)\beta(|x|)\mathrm{e}^{-2\pi i x\cdot\xi}\, dx = \sum_{\pm} \frac{1}{\pm 2\pi i|\xi|} \int_{1/4}^1 dr\ r^{d-1}\beta(r) a_\pm(r|\xi|) \frac{d}{dr}\mathrm{e}^{\pm 2\pi i r|\xi|} = \mathcal{O}(|\xi|^{-\frac{d+1}{2}})\,,$$

where the main contribution in the last step comes from the boundary term of the integration by parts. □

The other ingredient in the proof of (20.25) is

Theorem 20.18 (Poisson summation)**.** *If $\varphi \in \mathcal{S}(\mathbb{R}^d)$, then*

$$\sum_{j\in\mathbb{Z}^d} \varphi(j) = \sum_{j\in\mathbb{Z}^d} \hat{\varphi}(j)\,.$$

Proof. See, e.g., Grafakos [96, Theorem 3.2.8]. □

Proof of (20.25). If $\chi(x) \equiv \mathbf{1}_{B_0(1)}(x)$, then we can rewrite the assertion (20.25) as

$$N(\lambda) = \sum_{j\in\mathbb{Z}^d} \chi(j/\lambda) = |B_0(1)|\lambda^d + \mathcal{O}(\lambda^{d-2+\frac{2}{d+1}}), \quad \lambda \geq 1. \tag{20.27}$$

To prove this, we replace $\chi(x)$ by a smoother function that can be controlled using the Fourier transform and Poisson summation. To do so, fix $\beta \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$\beta \geq 0, \quad \int_{\mathbb{R}^d} \beta(y)\,dy = 1, \quad \text{and} \quad \beta(y) = 0, \quad \text{for } |y| \geq 1/2.$$

Then, for some $\varepsilon > 0$, depending on λ and to be specified later, we shall compare the sum in (20.27) to the smoothened version

$$\tilde{N}(\varepsilon,\lambda) := \sum_{j\in\mathbb{Z}^d} \tilde{\chi}_\lambda(\varepsilon, j), \tag{20.28}$$

where

$$\tilde{\chi}_\lambda(\varepsilon,x) := \left(\varepsilon^{-d}\beta(\cdot/\varepsilon) * \chi(\cdot/\lambda)\right)(x) = \int_{\mathbb{R}^d} \varepsilon^{-d}\beta\left((x-y)/\varepsilon\right)\chi(y/\lambda)\,dy.$$

Note that $0 \leq \tilde{\chi}_\lambda$, and, by the support properties of β, we also have $\chi(x/\lambda) = \tilde{\chi}_\lambda(\varepsilon,x)$ whenever $|x| \notin [\lambda-\varepsilon, \lambda+\varepsilon]$. Therefore,

$$\tilde{\chi}_{\lambda-\varepsilon}(\varepsilon,x) \leq \chi(x/\lambda) \leq \tilde{\chi}_{\lambda+\varepsilon}(\varepsilon,x),$$

i.e.,

$$\tilde{N}(\varepsilon,\lambda-\varepsilon) \leq N(\lambda) \leq \tilde{N}(\varepsilon,\lambda+\varepsilon). \tag{20.29}$$

Since $x \mapsto \tilde{\chi}_\lambda(\varepsilon,x)$ is Schwartz with Fourier transform given by

$$\lambda^d\hat{\chi}(\lambda\xi)\hat{\beta}(\varepsilon\xi),$$

Poisson summation gives (recalling $\int\chi = |B_0(1)|$ and $\int\beta = 1$)

$$\tilde{N}(\varepsilon,\lambda) = \lambda^d \sum_{j\in\mathbb{Z}^d} \hat{\chi}(\lambda j)\hat{\beta}(\varepsilon j) = |B_0(1)|\lambda^d + \lambda^d \sum_{\{j\in\mathbb{Z}^d:\, j\neq 0\}} \hat{\chi}(\lambda j)\hat{\beta}(\varepsilon j). \tag{20.30}$$

Since $|\hat{\beta}(\xi)| \lesssim_N (1+|\xi|)^{-N}$ (for any $N\in\mathbb{N}$) and $|\hat{\chi}(\xi)| \lesssim (1+|\xi|)^{-\frac{d+1}{2}}$ (by Corollary 20.17), the second term in (20.30) is bounded by

$$\lambda^d \sum_{\{j\in\mathbb{Z}^d:\, j\neq 0\}} (1+|\lambda j|)^{-\frac{d+1}{2}}(1+|\varepsilon j|)^{-N} \sim \lambda^d \int_{|\xi|\geq 1} (1+|\lambda\xi|)^{-\frac{d+1}{2}}(1+|\varepsilon\xi|)^{-N}\,d\xi.$$

for any $N\in\mathbb{N}$. But since for $0<\varepsilon\leq 1$ and $N>d$ one has

$$\begin{aligned}
&\int_{|\xi|\geq 1} (1+|\lambda\xi|)^{-\frac{d+1}{2}}(1+|\varepsilon\xi|)^{-N}\,d\xi \\
&\lesssim \int_{1\leq|\xi|\leq\varepsilon^{-1}} (1+|\lambda\xi|)^{-\frac{d+1}{2}}\,d\xi + \int_{|\xi|\geq\varepsilon^{-1}} (1+|\lambda\xi|)^{-\frac{d+1}{2}}(1+|\varepsilon\xi|)^{-N}\,d\xi \\
&\lesssim \lambda^{-\frac{d+1}{2}}\varepsilon^{-\frac{d-1}{2}} + (\lambda/\varepsilon)^{-\frac{d+1}{2}}\varepsilon^{-d} = \lambda^{-\frac{d+1}{2}}\varepsilon^{-\frac{d-1}{2}},
\end{aligned}$$

one concludes

$$\tilde{N}(\varepsilon,\lambda) = |B_0(1)|\lambda^d + \mathcal{O}(\lambda^{\frac{d-1}{2}}\varepsilon^{-\frac{d-1}{2}}).$$

Combining this with (20.29) thus yields

$$N(\lambda) = |B_0(1)|\lambda^d + \mathcal{O}(\varepsilon\lambda^{d-1}) + \mathcal{O}(\lambda^{\frac{d-1}{2}}\varepsilon^{-\frac{d-1}{2}}),$$

since $(\lambda \pm \varepsilon)^d = \lambda^d + \mathcal{O}(\varepsilon\lambda^{d-1})$ (coming from the $|B_0(1)|$ term). Optimizing in ε (i.e., choosing $\varepsilon = \lambda^{-\frac{d-1}{d+1}}$ so that both remainders are of the same order), finally shows the asserted (20.27). □

20.6. **Pseudodifferential operators.**

20.6.1. *Basics from the calculus of pseudodifferential operators.*

20.6.2. *Microlocal properties.* We shall now go over various microlocal properties of ΨDOs that we shall need later on. Among others, we shall give an equivalent definition of wave front sets that will be useful later on.

First, it will be useful to have microlocal versions of the existence of parametrices (i.e., approximate inverses) for elliptic pseudodifferential operators (satisfying $|P(x,\xi)| \geq c|\xi|^m$ for some $c > 0$, $m > 0$, and sufficiently large $|\xi|$). Recall that a parametrix of an elliptic ΨDO of order m is another ΨDO of order $-m$, say $E(x,D)$ having the property that, modulo smoothing operators in $S^{-\infty}$,

$$P \circ E = E \circ P = 1.$$

Any other operator with this property differs from E only via a smoothing operator. (See, e.g., Sogge [160, Theorem 4.2.5].)

To state a microlocal version of this fact [22], we need to denote the characteristic set of a ΨDO $P(x,D)$ of order m, which is a subset of $\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\} = T^*\mathbb{R}^d \setminus \{0\}$.

Definition 20.19. Let $P(x,D)$ be an elliptic ΨDO of order m. Then $\mathrm{Char}\, P$, the *characteristic set of* $P(x,D)$, is that closed subset of $T^*\mathbb{R}^d \setminus \{0\}$ whose complement is all points $(x_0,\xi_0) \in T^*\mathbb{R}^d \setminus \{0\}$ for which there is a conic neighborhood $\mathcal{N}_{x_0,\xi_0} \subseteq T^*\mathbb{R}^d \setminus \{0\}$ of (x_0,ξ_0) on which lower bounds of the form

$$|P(x,\xi)| \geq c|\xi|^m$$

hold for large $|\xi|$ with $c > 0$ possibly depending on $\mathcal{N}_{x_0,\xi_0}$.

Remark 20.20. Alternatively (as is standard in the analysis of differential operators $P(x,D) = \sum_{|\alpha|\leq m} a_\alpha(x) D^\alpha$), one could have simply defined

$$\mathrm{Char}\, P := \{(x,\xi) \in T^*(\mathbb{R}^d) \setminus \{0\} :\, P_m(x,\xi) = 0\},$$

where $P_m(x,\xi)$ is the principal part defined by $P_m(x,\xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha$.

The following is the microlocal version of the existence of parametrices for ΨDOs which are elliptic only in certain directions. For a symbol $a(x,\xi)$ and a conic neighborhood $\mathcal{N}$, we write $a \in S^{-\infty}(\mathcal{N})$, whenever we have for any N, α, and β that

$$\left|\left(\frac{\partial}{\partial x}\right)^\alpha \left(\frac{\partial}{\partial \xi}\right)^\beta a(x,\xi)\right| \lesssim_{N,\alpha,\beta} (1+|\xi|)^{-N}, \quad \text{if } (x,\xi) \in \mathcal{N}.$$

Theorem 20.21. *Let $P(x,D)$ be a ΨDO of order m and assume that $(x_0,\xi_0) \in T^*\mathbb{R}^d \setminus \{0\}$ is noncharacteristic for P, i.e.,*

$$(x_0,\xi_0) \notin \mathrm{Char}(P).$$

Then there is a ΨDO $E(x,D)$ of order $-m$ so that

$$(P \circ E)(x,\xi) - 1,\ (E \circ P)(x,\xi) - 1 \in S^{-\infty}(\mathcal{N})$$

for some conic neighborhood $\mathcal{N}$ of (x_0,ξ_0).

[22] that distinguishes between directions where $P(x,D)$ is elliptic and where not

Next, let us recall that ΨDOs are in general (as opposed to differential operators) non-local. Nonetheless, there remain certain remnants of locality in the sense that ΨDOs leave the singular support invariant. This fact is called *pseudolocality*. It means that if $P \in S^m$, then

$$\operatorname{sing\,supp} P(x,D)u \subseteq \operatorname{sing\,supp} u\,, \quad u \in H^{-\infty}\,. \tag{20.31}$$

This just follows from the fact that the kernel of $P(x,D)$ is smooth away from the diagonal. Similar considerations lead to the stronger microlocal property of $P(x,D)$, namely

$$WF(P(x,D)u) \subseteq WF(u)\,, \quad u \in H^{-\infty}\,. \tag{20.32}$$

Later on we shall show the almost inverse inclusion

$$WF(u) \subseteq WF(P(x,D)u) \cup \operatorname{Char}(P)\,,$$

see, e.g., Hörmander [111, Theorem 8.3.1] or Sogge [160, Corollary 4.2.11].

Remark 20.22. A final side remark is that if P is an elliptic differential operator, i.e., $P_m(x,\xi) \neq 0$ in $T^*(\mathbb{R}^d) \setminus \{0\}$, then we indeed have the reverse inclusion, i.e.,

$$WF(Pu) = WF(u)\,, \quad u \in \mathcal{D}'(\mathbb{R}^d)$$

and in particular $\operatorname{sing\,supp} Pu = \operatorname{sing\,supp} u$ for $u \in \mathcal{D}'(\mathbb{R}^d)$ (Hörmander [111, Corollary 8.3.2]).

Another fundamental object in microlocal analysis of ΨDOs is the notion of essential support.

Definition 20.23. Let P be a ΨDO. Then the *essential support of* $P(x,\xi)$, denoted by $\operatorname{ess\,sup} P$ is that closed subset of $T^*(\mathbb{R}^d) \setminus \{0\}$ whose complement consists of points (x_0,ξ_0) having the property that $P(x,\xi) \in S^{-\infty}(\mathcal{N}_{x_0,\xi_0})$ for some conic neighborhood $\mathcal{N}_{x_0,\xi_0}$ of (x_0,ξ_0) in $T^*(\mathbb{R}^d) \setminus \{0\}$.

Thus, if $u \in H^{-\infty}(\mathbb{R}^d)$, we have

$$WF(P(x,D)u) \subseteq \operatorname{ess\,sup} P\,.$$

Our next goal is to give an alternative characterization of $WF(u)$ whenever $u \in H^{-\infty}$. First, we note also that, by the definition of $WF(u)$, the statement $(x_0,\xi_0) \notin WF(u)$ means that $P(x,D)u \in C^\infty$ for certain ΨDOs $P(x,\xi) \in S^0$ that are non-characteristic at (x_0,ξ_0). Specifically, let

$$C_c^\infty(\mathbb{R}^d) \ni \rho(x) = \begin{cases} 1 & \text{for } |x| < 1/2 \\ 0 & \text{for } |x| \geq 1 \end{cases}$$

and smooth in between, and,

$$C_c^\infty(\mathbb{R}^d) \ni \chi(\xi) = \begin{cases} 0 & \text{for } |\xi| \ll 1 \\ 1 & \text{for } |\xi| \gg 1 \end{cases}\,.$$

Let us furthermore set

$$Q_\delta(x,D)v(x) := \int d\xi\, \mathrm{e}^{2\pi i x\cdot\xi} \rho(|x-x_0|/\delta)\rho\left(\left(\frac{\xi}{|\xi|} - \frac{\xi_0}{|\xi_0|}\right)/\delta\right)\chi(\xi)\hat{v}(\xi)\,.$$

Then $(x_0,\xi_0) \notin WF(u)$ if and only if $Q_\delta^* u \in C^\infty$ when $\delta > 0$ is small. This is due to the fact that the Fourier transform of $Q_\delta^* u$ equals

$$\rho((\frac{\xi}{|\xi|} - \frac{\xi_0}{|\xi_0|})/\delta)\chi(\xi)[\rho((\cdot - x_0)/\delta)u]^\wedge(\xi)\,.$$

Based on this, one checks that $(x_0,\xi_0) \notin WF(u)$ if and only if $Q_\delta^*(x,D)u \in \mathcal{S}(\mathbb{R}^d)$ when $\delta > 0$ is sufficiently small.

Moreover, since $P(x,D) = Q_\delta^*(x,D)$ is also not characteristic at (x_0,ξ_0) and we let for $m \in \mathbb{R}$

$$\mathcal{R}_m(u) := \{P(x,\xi) \in S^m :\, P(x,D)u \in C^\infty\}\,, \quad u \in H^{-\infty} \tag{20.33}$$

denote the set of *regularizing operators* for a given $u \in H^{-\infty}$ [23], then by the above arguments,

$$(WF(u))^c \subseteq \bigcup_{P \in \mathcal{R}_0(u)} (\mathrm{Char}\, P)^c\,.$$

The main result here is, however, that we actually have equality and not only for $m = 0$, but for all $m \in \mathbb{R}$. This provides a useful equivalent definition of $WF(u)$ (Definition 20.12).

Theorem 20.24. *Let $u \in H^{-\infty}$ and $m \in \mathbb{R}$. Then*

$$WF(u) = \bigcap_{P \in \mathcal{R}_m(u)} \mathrm{Char}\, P\,.$$

In particular, we have for a given $P \in S^m$,

$$WF(u) \subseteq \mathrm{Char}\, P\,, \quad \textit{if } P(x,D)u \in C^\infty\,.$$

The following corollary provides a nice complement of the microlocal property (20.32) of ΨDOs.

Corollary 20.25. *If $P \in S^m$ and $u \in H^{-\infty}$, then*

$$WF(u) \subseteq WF(P(x,D)u) \cup \mathrm{Char}\, P\,. \tag{20.34}$$

In particular, if u solves $P(x,D)u = 0$, then $WF(u) \subseteq \mathrm{Char}\, P$.

Proof. We prove the equivalent assertion

$$(WF(Pu))^c \cap (\mathrm{Char}\, P)^c \subseteq (WF(u))^c\,.$$

If $(x_0,\xi_0) \notin WF(Pu)$, then, by Theorem 20.24, there must be a $Q \in S^0$ with $Q(x,D) \circ P(x,D)u \in C^\infty$ and $(x_0,\xi_0) \notin \mathrm{Char}\, Q$. If also $(x_0,\xi_0) \notin \mathrm{Char}\, P$, then, by the Kohn–Nirenberg formula (cf. [160, Theorem 4.2.2])

$$(P \circ Q)(x,\xi) \sim \sum_\alpha \frac{1}{\alpha!} D_\xi^\alpha P(x,\xi) \left(\frac{\partial}{\partial x}\right)^\alpha Q(x,\xi)\,,$$

we also have $(x_0,\xi_0) \notin \mathrm{Char}(Q \circ P)$ and so $(x_0,\xi_0) \notin WF(u)$ also by Theorem 20.24. □

The proof of Theorem 20.24 relies on the following lemma, which is more or less equivalent to the theorem.

Lemma 20.26. *Let $u \in H^{-\infty}$. Then $(x_0,\xi_0) \notin WF(u)$ if and only if there is a conic neighborhood $\mathcal{N}$ of (x_0,ξ_0) in $T^*\mathbb{R}^d \setminus \{0\}$ so that $P(x,D)u \in C^\infty$ whenever $P(x,D)$ is a ΨDO with symbol $P(x,\xi)$ supported in $\mathcal{N}$.*

Let us first see how the Lemma implies the above theorem.

Proof of Theorem 20.24. Let $(x_0,\xi_0) \notin WF(u)$. Then by the lemma if $Q(x,\xi) \in S^m$ is supported in a small conic neighborhood of (x_0,ξ_0) and equals $|\xi|^m$ for $|\xi| \geq 1$ with $|\xi/|\xi| - \xi_0/|\xi_0||$ and $|x - x_0|$ small then $Qu \in C^\infty$. Since $Q(x,\xi)$ is non-characteristic at (x_0,ξ_0), we conclude

$$(x_0,\xi_0) \in \bigcup_{P \in \mathcal{R}_m(u)} (\mathrm{Char}\, P)^c\,,$$

[23] also including operators that may be characteristic at $(x_0,\xi_0) \notin WF(u)$

and thus

$$\bigcap_{P\in\mathcal{R}_m(R)} \operatorname{Char} P \subseteq WF(u)\,.$$

Conversely, suppose that $P(x,\xi)\in S^m$, $P(x,D)u\in C^\infty$, and $(x_0,\xi_0)\notin \operatorname{Char} P$. We then must show that $(x_0,\xi_0)\notin WF(u)$. By Theorem 20.21 we know that for such (x_0,ξ_0) there exists a microlocal parametrix $Q\in S^{-m}$ such that

$$(Q\circ P)(x,\xi)-1\in S^{-\infty}(\mathcal{N}_{x_0,\xi_0})$$

for some conic neighborhood $\mathcal{N}_{x_0,\xi_0}$ of (x_0,ξ_0). But then if $A\in S^\mu$ and $A(x,\xi)=0$ for $(x,\xi)\notin \mathcal{N}_{x_0,\xi_0}$ we have that $A(x,D)(Q\circ P-1)$ is smoothing by the Kohn–Nirenberg theorem (i.e., $P(x,D)\circ Q(x,D)$ is a ΨDO of order $m+\mu$ whenever P and Q are ΨDOs of order m, respectively μ, cf. [160, Theorem 4.2.2]). Since

$$u=Q(Pu)+(1-Q\circ P)u$$

and $Q(Pu)\in C^\infty$ (since $Pu\in C^\infty$), we conclude $A(x,D)u\in C^\infty$. Thus $(x_0,\xi_0)\notin WF(u)$ by the lemma which concludes the proof. □

Proof of Lemma 20.26. □

20.6.3. *Pseudodifferential operators on manifolds.* Before we define ΨDOs on manifolds and discussing some of their properties, we prove a preliminary result showing how certain types of ΨDOs on $\mathbb{R}^d$ transform under changes of coordinates.

We consider operators of the form

$$(P_\varphi u)(x)=\int \mathrm{e}^{2\pi i\varphi(x,y,\xi)}P(x,y,\xi)u(y)\,d\xi\,dy \tag{20.35}$$

where the compound symbol P belongs to S^m, i.e., satisfies

$$\left|\left(\frac{\partial}{\partial\xi}\right)^\alpha\left(\frac{\partial}{\partial x}\right)^{\beta_1}\left(\frac{\partial}{\partial y}\right)^{\beta_2}P(x,y,\xi)\right|\lesssim_{\alpha,\beta}(1+|\xi|)^{m-|\alpha|}$$

and $\varphi\in C^\infty(\mathbb{R}^{2d}\times\mathbb{R}^d\setminus\{0\})$ is real-valued, homogeneous of degree one in ξ, and satisfies

$$\varphi(x,x,\xi)\equiv 0 \quad\text{and}\quad \nabla_x\varphi(x,y,\xi)\big|_{x=y}\equiv\xi\,. \tag{20.36}$$

In particular this means that $\mathrm{e}^{i\varphi}$ behaves like a plane wave near the diagonal, i.e., one has

$$\varphi(x,y,\xi)=\langle(x-y),\xi\rangle+\mathcal{O}(|x-y|^2||\xi|)\,.$$

Thus, if $\operatorname{supp}_{x,y}P$ is contained in a sufficiently small neighborhood of the diagonal, we have that

$$|\nabla_\xi(\varphi(x,y,\xi)-\langle x-y,\xi\rangle|\le\frac12|x-y|\quad\text{on }\operatorname{supp}P\,. \tag{20.37}$$

Under these hypotheses, we have the following

Proposition 20.27. *Suppose $P\in S^m$ as above vanishes when x or y is outside of a fixed compact set in $\mathbb{R}^d$ and that φ satisfies* (20.36) *and* (20.37)[24]*. Then P_φ is a ΨDO of order m. Moreover, if we set $P(x,\xi)=P(x,x,\xi)$, then $P_\varphi-P(x,D)$ is a ΨDO of order $m-1$.*

Proof. See Sogge [160, Proposition 4.2.12]. □

[24] As we have seen above, the second condition is actually a consequence of the former, but we nevertheless include it in the statement for the sake of clarity.

We will now apply this result to see how ΨDOs in $\mathbb{R}^d$ behave under changes of variables. For simplicity, we assume for the moment that the operators have symbols satisfying $P(x,\xi)=0$ for x outside of a compact set K. Recall that if $\kappa:\mathbb{R}^d\to\mathbb{R}^d$ is a diffeomorphism, then the pullback of a function $u\in C^\infty$ via κ is $\kappa^* u = u_\kappa$, defined by

$$u_\kappa(x)=u(\kappa(x))\,.$$

Proposition 20.28. *Let $\kappa:\mathbb{R}^d\to\mathbb{R}^d$ be a diffeomorphism and assume that $P(y,\xi)\in S^m$ vanishes when y is outside of a compact set K. Then there is a symbol $P_\kappa(x,\xi)\in S^m$ such that, modulo smoothing operators,*

$$(P_\kappa(x,D)u_\kappa)(x)=(P(y,D)u)(y)\,,\quad y=\kappa(x)\,,$$

and

$$P_\kappa(x,\,{}^t\kappa'(x)\xi)-P(\kappa(x),\xi)\in S^{m-1}\,. \tag{20.38}$$

Remark 20.29. Note that (20.38) says that, modulo symbols of one order less, the symbols of ΨDOs pull back according to the pullback map

$$(\kappa(x),\xi)\mapsto(x,\kappa'(x)^t\xi)$$

which is the change of variables formula for the cotangent bundle coming from changes of coordinates in the base. This fact will momentarily tell us that the principal symbol of a ΨDO on a manifold M is invariantly defined as a function on $T^*M\setminus 0$.

Proof of Proposition 20.28. Choose $\rho\in C_c^\infty(\mathbb{R}^d)$ satisfying $\rho(y)=1$ near $y=0$. Then if we set $y=\kappa(x)$, $z=\kappa(w)$ and $\xi={}^t\kappa'(x)\eta$, we obtain, modulo a smoothing operator, that $P(y,D)$ is given by

$$\int \mathrm{e}^{2\pi i\langle y-z,\eta\rangle}P(y,\eta)\rho(z-y)u(z)\,d\eta\,dz=\int \mathrm{e}^{2\pi i\varphi(x,\omega,\xi)}Q(x,w,\xi)u_\kappa(w)\,d\xi\,dw\,,$$

where

$$\varphi(x,w,xi)=\langle\kappa(x)-\kappa(w),({}^t\kappa'(x))^{-1}\xi\rangle$$

and

$$Q(x,w,\xi)=\rho(\kappa(w)-\kappa(x))P(\kappa(x),({}^t\kappa'(x))^{-1}\xi)|\kappa'(w)||{}^t\kappa'(x)|^{-1}\,.$$

Since φ is as in (20.36) and since

$$Q(x,w,{}^t\kappa'(x)\eta)\big|_{w=x}=P(\kappa(x),\eta)\,,$$

the claim follows from Proposition 20.27. □

We may now define ΨDOs on a smooth compact manifold M.

Definition 20.30. A map $P:C^\infty(M)\to C^\infty(M)$ is called a *ΨDO of order m* if its kernel is smooth away from the diagonal $\Delta=\{(x,y)\in M\times M:\,x=y\}$, and, whenever $\Omega_\nu\subseteq M$ is a coordinate patch with coordinates

$$y=\kappa_\nu(x)\in\tilde\Omega_\nu:=\kappa_\nu(\Omega_\nu)\subseteq\mathbb{R}^d\,,\quad x\in\Omega_\nu\,,$$

and $\psi_\nu,\tilde\psi_\nu\in C_c^\infty(\tilde\Omega_\nu)$, the operators

$$P_\nu u(y)=\tilde\psi_\nu(\kappa_\nu(x))P\left((\psi_\nu u)\circ\kappa_\nu(\cdot)\right)(x)\,,\quad y=\kappa_\nu(x)\in\kappa_\nu(\Omega_\nu)\subseteq\mathbb{R}^d\,,\quad u\in C^\infty(\mathbb{R}^d) \tag{20.39}$$

are (usual) ΨDOs of order m.

In this formula $(\psi_\nu u)\circ\kappa_\nu$ is understood to be the $C^\infty(M)$ function which equals $\psi_\nu(\kappa_\nu(x))u(\kappa_\nu(x))$ when $\kappa_\nu(x)\in \operatorname{supp}\psi_\nu$ and zero otherwise. If $\bigcup_\nu \Omega_\nu = M$ is a finite covering of M by coordinate patches and $\{\Psi_\nu\}$ is a smooth partition of unity subordinate to this covering, i.e., $\sum_\nu \Psi_\nu \equiv 1$ and $\operatorname{supp}\Psi_\nu \subseteq \Omega_\nu$, and if $\tilde\Psi_\nu \in C^\infty(M)$ equals one on $\operatorname{supp}\Psi_\nu$ and is supported in Ω_ν for each ν, then, modulo an operator with smooth kernel (i.e., a smoothing operator), we have $Pv = \sum_\nu \tilde\Psi_\nu P(\Psi_\nu v)$. Consequently, we can use (20.39) with ψ_ν and $\tilde\psi_\nu$ being the pushforwards of Ψ_ν and $\tilde\Psi_\nu$ respectively, to write the symbol of P in local coordinates as a function $P(y,\eta) = P_\nu(y,\eta)\in S^m$.

Definition 20.31. We say that P is a *classical* ΨDO *of order* m and write $P\in\Psi^m_{cl}(M)$ if in every local coordinate system, we have

$$P(y,\eta)\sim\sum_{j=0}^\infty P_{m-j}(y,\eta)\,,$$

where P_{m-j} is homogeneous of degree $m-j$ in η.

We shall restrict ourselves to such polyhomogeneous operators from now on since operators such as $\sqrt{-\Delta_g}$ always have this form. As usual, Δ_g denotes the Laplace–Beltrami operator on M endowed with a Riemannian metric g.

If we use local coordinates (cf. Sogge [160, Section §2.3])

$$T^*M\ni(x,\xi)\mapsto(\kappa_\nu,\xi^\nu)\in\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}\,,\quad x\in\Omega_\nu\subseteq M\,,$$

then we can define the *principal part of a classical* ΨDO P by setting

$$p(x,\xi)=P_m(\kappa_\nu(x),\xi^\nu)\,.$$

By Remark 20.29, this gives a well-defined function on $C^\infty(T^*M\setminus 0)$ which is homogeneous of degree m. Naturally, we say that P *is elliptic* if its principal symbol never vanishes on $T^*M\setminus 0$. Moreover, we define the characteristic set of P as

$$\operatorname{Char}P=\{(x,\xi)\in T^*M\setminus 0:\, p(x,\xi)=0\}\,.$$

As we indicated in Remark 20.16, the wave front set of $V\in H^{-\infty}(\mathbb{R}^d)$ transforms according to the change of variables formula (a.k.a. the pullback formula for the cotangent bundle)

$$(\kappa(x),\xi)\mapsto(x,\kappa'(x)^t\xi)$$

for the cotangent bundle. That means that we can use local coordinates to define the wave front set of a given $u\in H^{-\infty}(M)=\bigcup_s H^s(M)$. For such u it is also clear that, if $\mathcal{R}_m(u)$ denotes those $P\in\Psi^m_{cl}(M)$ for which $Pu\in C^\infty(M)$, then, by Theorem 20.24, for each $m\in\mathbb{R}$ we have

$$WF(u)=\bigcap_{P\in\mathcal{R}_m(u)}\operatorname{Char}P\,. \tag{20.40}$$

In particular, by using local coordinates, we see that the notion of essential support of $P(x,\xi)\in\Psi^m_{\mathcal{L}}(M)$ is a well-defined subset of $T^*M\setminus 0$, and so, as in the euclidean case (Theorem 20.24) we have

$$WF(P(x,D)u)\subseteq\operatorname{ess\,supp}P\,,\quad u\in H^{-\infty}(M)\,.$$

If we are working on a Riemannian manifold (M,g), then $P\in\Psi^m_{cl}(M)$ is said to be *self-adjoint* if

$$(Pu,v)=(u,Pv):=\int_M\overline{u}Pv\,dV_g\,,\quad u,v\in C^\infty(M)\,.$$

Recall that $P(x,\xi)-\mathrm{Re}(P(x,\xi)) \in S^{m-1}$ for m-th order, self-adjoint ΨDOs P (cf. [160, Corollary 4.2.8]). Thus, if $P \in \Psi_{cl}^m(M)$ is self-adjoint and elliptic, then its principal symbol must be real and either be always positive or always negative on $T^*M \setminus 0$.

As usual, we can define Sobolev saces of order s on M by setting

$$\|f\|_{H^s(M)} = \sum_\nu \|f_\nu\|_{H^s(\mathbb{R}^d)}\,, \quad f_\nu(y) = (\Psi_\nu f)(x)\,, \quad y = \kappa_\nu(x)\,, \quad x \in \operatorname{supp}\Psi_\nu$$

where, as before $\{\Psi_\nu\}$ is a smooth partition of unity coming from a finite covering of M by the coordinate patches (Ω_ν, κ_ν). It is straightforward to check that different partitions of unity give comparable Sobolev norms. Thus, there is no loss in just defining the Sobolev norms via one of them. Moreover, in view of classical ΨDO calculus, we have

$$P : H^s(M) \to H^{s-m}(M)\,, \quad P \in \Psi_{cl}^m(M)\,.$$

If $m > 0$ and $P \in \Psi_{cl}^m(M)$ is elliptic, then

$$\|u\|_{H^m} \lesssim \|Pu\|_{L^2(M)} + \|u\|_{L^2(M)}\,.$$

If $m = 1$ and $Q \in \Psi_{cl}^1(M)$ is self-adjoint and elliptic, then, as noted above, after possibly multiplying by -1, we may assume its principal symbol, $q(x,\xi)$, to be positive. Then if $A \in \Psi_{cl}^{1/2}(M)$ has principal symbol $\sqrt{q(x,\xi)}$, the previous inequality shows

$$\|u\|_{H^{1/2}(M)} \lesssim \|Au\|_{L^2(M)}^2 + \|u\|_{L^2(M)}^2\,.$$

Since $Q - A^*A \in \Psi_{cl}^0(M)$, Cauchy–Schwarz gives

$$|(u,Qu) - (u,A^*Au)| \lesssim \|u\|_{L^2(M)}^2$$

and therefore, by combining the last two inequalities and noting $(u, A^*Au) = \|Au\|^2$,

$$\|u\|_{H^{1/2}(M)}^2 \lesssim (u,Qu) + \|u\|^2 = (u,(Q+1)u)\,.$$

Thus, $Q+c$ is a positive self-adjoint operator, and, by Rellich–Kondrachov, has compact resolvent, so purely discrete spectrum consisting of eigenvalues $0 < \mu_1 \le \mu_2 \le ...\mu_j$ possibly accumulating at infinity. In particular, Q has also purely discrete spectrum, possibly accumulating at $+\infty$ with only finitely many negative eigenvalues (if any).

We are now prepared to study $\sqrt{-\Delta_g}$ and show that it belongs to $\Psi_{cl}^1(M)$ with principal symbol

$$p_{\sqrt{-\Delta_g}}(x,\xi) = \sqrt{\sum_{j,k=1}^d g^{jk}(x)\xi_j\xi_k} \tag{20.41}$$

whenever $-\sum_{j,k=1}^d g^{jk}(x)\xi_j\xi_k$ is the principal symbol of the Laplace–Beltrami operator Δ_g on our Riemannian manifold (M,g).

Recall that if $0 = \lambda_0^2 < \lambda_1^2 \le \lambda_2^2 \le ...$ are the eigenvalues of $-\Delta_g$ with corresponding eigenprojections E_j, then

$$-\Delta_g u = \sum_{j\ge 0} \lambda_j^2 E_j u\,, \quad u \in C^\infty(M)$$

so naturally (by functional calculus), we define $P = \sqrt{-\Delta_g}$ by

$$Pu = \sum_{j\ge 0} \lambda_j E_j u\,, \quad u \in C^\infty$$

which satisfies $P^2 = P \circ P = -\Delta_g$.

Because of the zero-eigenvalue λ_0, the operator $-\Delta_g$ is not invertible. However, by modifying it by the rank-one projection E_0, i.e., setting

$$Lu = E_0 u + \sum_{j\geq 1} \lambda_j^2 E_j u\,, \quad u \in C^\infty(M)\,,$$

we see that $L > 0$ is invertible and that L only differs from $-\Delta_g$ by E_0, i.e., a smoothing operator with kernel $(\mathrm{vol}_g(M))^{-1}$ on $M \times M$.

We now show that $P \in \Psi^1_{cl}(M)$. To do so, we first construct a positive first order self-adjoint, elliptic operator $Q \in \Psi^1_{cl}(M)$ that satisfies

$$L - Q^2 = R \tag{20.42}$$

for some smoothing operator R. Working in local coordinates, we first set

$$\tilde{Q}_1(x,\xi) = \chi(\xi)\left(\sum_{j,k=1}^d g^{jk}(x)\xi_j\xi_k\right)^{1/2}\,,$$

where $\chi \in C^\infty$ vanishes near 0 but equals 1 when, say, $|\xi||geq1$. If we let $Q_1(x,D) = (\tilde{Q}_1(x,D) + \tilde{Q}_1^*(x,D))/2$, then Q_1 is self-adjoint, and in $\Psi^1_{cl}(M)$. Moreover (by the Kohn–Nirenberg theorem [160, Theorem 4.2.2]) $(Q_1)^2 - L \in \Psi^1_{cl}(M)$. We can now continue inductively choosing self-adjoint $Q_j \in \Psi^{2-j}_{cl}(M)$ $(j = 2, 3.,,,)$ so that $L - (Q_1^2 + \cdots + Q_N)^2 \in \Psi^{2-N}_{cl}(M)$. As a result, if we let $Q \in \Psi^1_{cl}(M)$ be a representative of the formal series $\sum_{j\geq 1} Q_j$, we would get that $L - Q^2$ is smoothing. Since each Q_j is self-adjoint, Q equals its adjoint by a smoothing error. Thus, after possibly adding such a smoothing error operator, we may indeed assume Q to be self-adjoint. By what we did before, Q then has discrete spectrum accumulating at $+\infty$. Thus, after possibly modifying it on a one- (or finite-) dimensional set, we may also assume that Q is positive and that $L - Q^2 = R$ indeed holds, as claimed.

Summarized, we found an approximation, i.e., Q^2, of $L = -\Delta_g + E_0$. We now claim that also $\sqrt{L} - Q \equiv R_0$ is smoothing. To see this let $\gamma \subseteq \mathbb{C}$ be a contour encircling all eigenvalues of L. Then by Cauchy's integral formula,

$$L^{-1/2} = -\frac{1}{2\pi i}\int_\gamma z^{-1/2}(L - z)^{-1}\,dz$$

and

$$Q^{-1} = -\frac{1}{2\pi i}\int_\gamma z^{-1/2}(Q^2 - z)^{-1}\,dz = -\frac{1}{2\pi i}\int_\gamma z^{-1/2}(L - R - z)^{-1}\,dz$$

and therefore,

$$\begin{aligned} L^{-1/2} - Q^{-1} &= -\frac{1}{2\pi i}\int_\gamma z^{-1/2}\left[(L-z)^{-1} - (L-R-z)^{-1}\right]\,dz \\ &= \frac{1}{2\pi}\int_\gamma z^{-1/2}\left[(L-z)^{-1}R(L-R-z)^{-1}\right]\,dz\,. \end{aligned}$$

Since R is smoothing, the whole integrand is smoothing and the integral in particular converges and defines a smoothing operator. Thus

$$\sqrt{L} - Q = \sqrt{L} - \sqrt{-\Delta_g} + \sqrt{-\Delta_g} - Q = Q(Q^{-1} - L^{-1/2})L^{1/2} \equiv R_0$$

is then smoothing as well, we obtain the claim. Since $\sqrt{L} - \sqrt{-\Delta_g}$ is a rank-one projection onto constant functions, it follows from $\sqrt{L} - Q$ being smoothing that $\sqrt{-\Delta_g} - Q$ is smoothing, too. In summary, we have proven

Theorem 20.32. *Let Δ_g be the Laplace–Beltrami operator on a compact Riemannian manifold (M,g). Then $P := \sqrt{-\Delta_g} \in \Psi^1_{cl}(M)$ is a self-adjoint, first-order classical ΨDO with principal symbol $p(x,\xi) = \sqrt{\sum_{j=k=1}^d g^{jk}(x)\xi_j\xi_k}$.*

Similar arguments show that the operators defined by

$$(1-\Delta_g)^{s/2} f = \sum_{j\geq 0} (1+\lambda_j^2)^{s/2} E_j f\,, \quad f \in C^\infty(M)\,, \quad s \in \mathbb{R}$$

belong to $\Psi^m_{cl}(M)$ with principal symbol

$$\left(1 + \sum_{j,k=1}^d g^{jk}(x)\xi_j\xi_k\right)^{s/2}.$$

Moreover, for each $s \in \mathbb{R}$ we have

$$\|u\|_{H^s(M)} \sim \|(1-\Delta_g)^{s/2} u\|_{L^2(M)}$$

which just follows from $(1-\Delta_g)^{s/2} : H^s \to L^2$ and $(1-\Delta_g)^{-s/2} : L^2 \to H^s$ boundedly.

20.7. **Propagation of singularities and Egorov's theorem.** We follow Sogge [160, Section 4.3].

Throughout this section we always take

$$P = \sqrt{-\Delta_g}$$

and are concerned with the associated Schrödinger (or in this case, the half-wave) equation

$$\begin{cases} (\partial_t - iP(x,D))u(t,x) = F(x,t)\,, \quad 0 < t < T \\ u\big|_{t=0} = f\,. \end{cases} \tag{20.43}$$

Clearly, its solution is given by the Duhamel formula

$$u(x,t) = (\mathrm{e}^{itP})f(x) + i\int_0^t \mathrm{e}^{i(t-s)P} F(s,x)\,ds\,.$$

Before we discuss the solution operator e^{itP} in more detail, let us go over some basic properties of the solution directly via energy estimates. The following lemma resembles that for the usual wave equation, cf. Sogge [160, Formula (3.1.17)].

Lemma 20.33. *Let $s \in \mathbb{R}$. If*

$$u \in C^1([0,T] : H^s) \cap C([0,T] : H^{s+1})\,,$$

then there is a constant C_s, independent of T such that

$$\sup_{t\in[0,T]} \|u(t,\cdot)\|_{H^s(M)} \leq C_s \left(\|u(0,\cdot)\|_{H^s(M)} + \int_0^T \|(\partial_t - iP)u(t,\cdot)\|_{H^s(M)}\,dt \right). \tag{20.44}$$

Proof. □

These energy estimates allow one to prove an existence and uniqueness theorem for the half-wave equation (20.43).

Theorem 20.34. *Let $s \in \mathbb{R}$. Then for every $F \in L^1([0,T] : H^s)$ and $f \in H^s$ there is a unique solution $u \in C([0,T] : H^s)$ of the Cauchy problem* (20.43) *and it must satisfy* (20.44).

Proof. □

This result gives the following

Corollary 20.35. *Let $F \equiv 0$ and suppose u satisfies (20.43) with $f \in H^s$ for every $s \in \mathbb{R}$. Then if $u \in C^1([0,T] : H^{s_0}(M))$ for some $s_0 \in \mathbb{R}$, it follows that $u \in C^1([0,T] : H^s(M))$ for every s and the same is true for $\partial_t^j u$ for any $j \in \mathbb{N}$. Thus, $u \in C^\infty(\mathbb{R} \times M)$.*

The main interest of this section is the propagation of singularities for the half-wave equation (20.43). The analysis relies on the following

Proposition 20.36. *Let $Q \in \Psi^m_{cl}(M)$. Then there exists a one-parameter family of ΨDOs $t \mapsto E(t) \in \Psi^m_{cl}(M)$ depending smoothly on t and satisfying*

$$[\partial_t - iP, E(t)] = 0\,, \quad E(0) = Q\,, \tag{20.45}$$

and having for each $t \in \mathbb{R}$ the principal symbol

$$E_0(t; x, \xi) = q_0(\Phi_t(x, \xi)) \tag{20.46}$$

with $q_0(x, \xi)$ being the principal symbol of Q and where $\Phi_t : T^*M \setminus 0 \to T^*M \setminus 0$ being the Hamiltonian flow for to the Hamiltonian vector field

$$H_p := \frac{\partial p}{\partial \xi}\frac{\partial}{\partial x} - \frac{\partial p}{\partial x}\frac{\partial}{\partial \xi} \tag{20.47}$$

associated to the principal symbol $p(x, \xi)$ of P (cf. (20.41)).

Before turning to the proof, we state some immediate consequences thereof. The first concerns Hörmander's theorem about propagation of singularities of solutions to (20.43).

Theorem 20.37. *Let $f \in H^{-\infty}(M)$ and let $u \in C([0,T] : H^{-\infty}(M))$ be the solution of the homogeneous Cauchy problem*

$$(\partial_t - iP)u = 0\,, \quad u\big|_{t=0} = f\,. \tag{20.48}$$

Then for each fixed $t \in \mathbb{R}$, we have $\Phi_t(WF(u(t,\cdot))) = WF(f)$, i.e.,

$$WF(u(t,\cdot)) = \{(y, \eta) \in T^*M \setminus 0 : \Phi_t(y, \eta) = (x, \xi) \text{ for some } (x, \xi) \in WF(f)\}\,. \tag{20.49}$$

Besides the above propagation of singularities result, we have the following special case of *Egorov's theorem* as a consequence of Proposition 20.36.

Theorem 20.38 (Egorov (special case))**.** *If $Q \in \Psi^m_{cl}(M)$ with principal symbol $q_0(x, \xi)$ then*

$$e^{itP} Q e^{-itP} \tag{20.50}$$

is a one-parameter family of ΨDOs $E_Q(t) \in \Psi^m_{cl}(M)$ depending smoothly on $t \in \mathbb{R}$. Their principal symbol is given by $q_0(\Phi_t(x, \xi))$ where Φ_t is the Hamiltonian flow associated to the principal symbol of $P = \sqrt{-\Delta_g}$.

Proof. □

Remark 20.39. One can easily prove that the principal symbol of

$$Q(t; x, D) = e^{itP} Q(x, D) e^{-itP}$$

is $q_0(\Phi_t(x, \xi))$ if one just assumes that the evolved $Q(t; x, D)$ is a ΨDO. The latter in turn can be verified for small $|t|$ using the Hadamard parametrix (cf. Sogge [160, Theorem 2.4.1]), Theorem 20.32 (on the fact that $\sqrt{-\Delta_g} \in \Psi^1_{cl}(M)$), and the proof of [160, Lemma 5.2.2]. Once the small $|t|$ result is established, the large result continues to hold for all $t \in \mathbb{R}$ by iteration using the group property

$$e^{i(t_1+t_2)P} = e^{it_1 P} e^{it_2 P}\,.$$

Now we verify the initial claim assuming $Q(t) \equiv Q(t; x, D)$ is a ΨDO with principal symbol $q_0(t)$. First note

$$\partial_t Q(t) = i[P, Q(t)]$$

and recall that the commutator of two ΨDOSs is of one order lower than their sum and that its symbol is given by the Poisson bracket of their symbols (cf. [160, Corollary 4.2.3]). Thus, the principal symbol $\partial_t q_0(t)$ of $\partial_t Q(t) = i[P, Q(t)]$ is given by

$$\partial_t q_0(t) = \{p, q_0(t)\} = H_p q_0(t) = \frac{\partial p}{\partial \xi} \cdot \frac{\partial q_0(t)}{\partial x} - \frac{\partial p}{\partial x} \cdot \frac{\partial q_0(t)}{\partial \xi} \, .$$

This equation has a unique solution which satisfies the initial condition

$$q_0(0; x, \xi) = q_0(x, \xi) \, .$$

Since $q_0(\Phi_t(x,\xi))\big|_{t=0} = q_0(x,\xi)$ and (by the classical Hamiltonian equations of motion $\Phi_t(x,\xi) = (x(t), \xi(t))$ with $\dot{x}(t) = \partial_\xi p$ and $\dot{\xi}(t) = -\partial_x p$)

$$\begin{aligned}
\partial_t q_0(\Phi_t(x,\xi)) &= \frac{\partial q_0(\Phi_t(x,\xi))}{\partial x} \cdot \frac{dx(t)}{dt} + \frac{\partial q_0(\Phi_t(x,\xi))}{\partial x} \cdot \frac{d\xi(t)}{dt} \\
&= \frac{\partial q_0(\Phi_t(x,\xi))}{\partial x} \cdot \frac{\partial p}{\partial \xi} - \frac{\partial q_0(\Phi_t(x,\xi))}{\partial \xi} \cdot \frac{\partial p}{\partial x} = H_p q_0(\Phi_t(x,\xi)) \, ,
\end{aligned}$$

we indeed conclude $q_0(t; x, \xi) = q_0(\Phi_t(x,\xi))$.

20.8. **Friedrichs' quantization and the sharp Gårding inequality.** A procedure that assigns to a function $P(x,\xi) \in S^m$ (a symbol) an operator on $H^{-\infty}(\mathbb{R}^d)$ is called a quantization. The *Kohn–Nirenberg quantization*

$$P(x, D)u = \int \mathrm{e}^{2\pi i (x-y)\cdot\xi} P(x,\xi) u(y) \, dy \, d\xi = \int \mathrm{e}^{2\pi i x\cdot\xi} P(x,\xi) \hat{u}(\xi) \, , \quad u \in \mathcal{S}(\mathbb{R}^d) \tag{20.51}$$

is simple and natural as it closely resembles Fourier multiplier operators.

In application to quantum mechanics one would like the quantization of $P(x,\xi)$ to be self-adjoint if the symbol is real. However, this is not the case for the Kohn–Nirenberg quantization but at least for the *Weyl quantization*

$$P_W u(x) = \int \mathrm{e}^{2\pi i (x-y)\cdot\xi} P\left(\frac{x+y}{2}, \xi\right) u(y) \, dy \, d\xi \, . \tag{20.52}$$

We will not make use of this quantization but excellent references describing it and the resulting calculus include Martinez [135, Theorem 2.7.1] (showing how to change between different quantizations), as well as Folland [78], Muscalu–Schlag [141], and Hörmander [110].

Another desirable feature – that is lacking in the Kohn–Nirenberg quantization too – is that the quantized operators are non-negative whenever their symbols are. The Friedrichs quantization [89] that we are about to discuss now remedies this failure. It is particularly useful in the study of quantum ergodicity, see also Sogge [160, Chapter 6].

Example 20.40. The following example illustrates that the Kohn–Nirenberg quantization does not preserve non-negativity. Consider, e.g., $a(x,\xi) := a(x)\xi^2$ with $0 \leq a(x) \in C_c^\infty(\mathbb{R})$. Then the associated operator $-a(x)d_x^2$ is in general not non-negative. For instance, if $u \in C_c^\infty(\mathbb{R})$ is such that $u(x) = u''(x)$ for all $x \in \mathrm{supp}(a(x))$, then $(u, a(x,D)u) = -\int_{\mathbb{R}} |u(x)|^2 a(x) < 0$. On the other hand, the operator $-d_x a(x) d_x$ is indeed non-negative and it agrees with the Kohn–Nirenberg quantizaed $a(x,D)$ up to an operator of lower order.

We will now consider a similar construction for general ΨDOs. Specifically, we show that if $0 \leq a(x,\xi) \in S^m$, then, up to an operator of one order less, $a(x,D)$ is also nonnegative.

Theorem 20.41 (Friedrichs)**.** *Let $a \in S^\mu$ and assume that $a(x,\xi) \geq 0$. Then one can write*

$$a(x,\xi) = a_F(x,\xi) + r(x,\xi) \tag{20.53}$$

where $r \in S^{\mu-1}$ and

$$(u, a_F(x,D)u) \geq 0\,, \quad u \in \mathcal{S}\,. \tag{20.54}$$

In particular, one choice for such an $a_F(x,\xi)$ is

$$a_F(x,\xi) = \int \psi\left((x-y)q(\eta), (\xi-\eta)/q(\eta)\right) a(y,\eta)\, dy\, d\eta \tag{20.55}$$

where $q(\eta) = (1+|\eta|^2)^{1/4}$ and $\psi(x,\xi) \in \mathcal{S}(\mathbb{R}^{2d})$ is the integral kernel of $\psi(x,D) = \varphi(x,D)^*\varphi(x,D)$ where $\varphi \in C_c^\infty(\mathbb{R}^{2d})$ is even with $\|\varphi\|_2 = 1$.

Proof. See Sogge [160, Theorem 4.4.1]. □

Importantly, this result (and $\|u\|_{H^m} \lesssim \|Pu\|_2 + \|u\|_2$ for any ΨDO P of order m) immediately gives

Corollary 20.42 (Sharp Gårding inequality)**.** *If $a \in S^{2m+1}$ and $\mathrm{Re}(a(x,\xi)) \geq 0$, then*

$$\mathrm{Re}(u, a(x,D)u) \gtrsim -\|u\|_{H^m}^2\,, \quad u \in \mathcal{S}\,. \tag{20.56}$$

Proof. We write

$$(\mathrm{Re}\, a)(x,D) = \frac{a(x,D) + a(x,D)^*}{2} + \left((\mathrm{Re}\, a)(x,D) - \frac{a(x,D)+a(x,D)^*}{2}\right)$$

and notice that the term in parantheses is a ΨDO of order $2m$. Since $\|u\|_{H^m} \lesssim \|Pu\|_2 + \|u\|_2$ for any ΨDO P of order m, it suffices to prove the assertion for $a(x,D)$ instead of $(\mathrm{Re}\, a)(x,D)$. Thus, we can without loss of generality assume $a(x,\xi) \geq 0$. But now we can apply Friedrichs' theorem and are done since $r \in S^{2m}$. □

The following generalizes Theorem 20.41 to Riemannian manifolds.

Theorem 20.43. *Let (M,g) be a Riemannian manifold of dimension d. Then there is a linear map*

$$a(x,\xi) \mapsto a_F(x,D)$$

sending each function $a \in C^\infty(T^*M \setminus 0)$ which is homogeneous of degree zero in ξ to a ΨDO $a_F(x,D)$ such that the principal part of $a_F(x,D)$ equals $a(x,\xi)$ and, moreover,

$$(h, a_F(x,D)h) \geq 0\,, \quad h \in L^2(M)\,, \quad \text{if } a(x,\xi) \geq 0\,. \tag{20.57}$$

Moreover, if $A(x,D) \in \psi^0_{cl}(M)$ is a classical ΨDO with principal symbol $a(x,\xi)$, then $a_F(x,D) - A(x,D)$ is of order -1.

Proof. After a partition of unity involving non-negative functions, we may assume that $a(x,\xi)$ vanishes when x is outside of a compact subset of a coordinate patch. We may also supppose that the support of $a(x,\xi)$ is so small that coordinates can be chosen so that $|h| \equiv 1$ in the coordinate patch. If we then work in local coordinates and let $\tilde{a}_F(x,\xi)$ denote the right side of (20.55), we obtain a ΨDO $\tilde{a}_F(x,D)$ with principal symbol $a(x,\xi)$ which is non-negative on $L^2(\mathbb{R}^d)$ if $a(x,\xi) \geq 0$. If $0 \leq \varphi \in C^\infty(\mathbb{R}^d)$ and $\varphi(x) = 1$ on the x-support of a, then the same is true for the operator $\varphi\tilde{a}_F(x,D)\varphi$. If we assume as well that φ is supported in the image of our coordinate patch, then the pullback, i.e., $a_F(x,D)$, of this operator to M will have the desired properties. □

Let us finally denote by $p(x,\xi)$ the principal symbol of $\sqrt{-\Delta_g}$ and define its unit cotangent bundle

$$S^*M = \{(x,\xi) \in T^*M : p(x,\xi) = 1\}.$$

Then to every $a_0 \in C^\infty(S^*M)$ we can naturally associate a homogeneous of degree zero function (extension) $a(x,\xi) \in C^\infty(T^*M \setminus 0)$ given by $a(x,\xi) = a_0(x, \xi/p(x,\xi))$. Then, using Theorem 20.43 we can easily obtain the following result saying that (20.57) associates to each $h \in L^2(M)$ a natural distribution on S^*M.

Corollary 20.44. *Let (M,g) be as above and fix $h \in L^2(M)$. Also given $a_0 \in C^\infty(S^*M)$ as above, let $a \in C^\infty(T^*M \setminus 0)$ denote its homogeneous of degree zero extension and a_F its corresponding Friedrichs quantization in Theorem 20.43. Then the map*

$$C^\infty(S^*M) \ni a_0 \mapsto u_h(a_0) = (h, a_F(x,D)h)$$

defines a non-negative distribution $u_h \in \mathcal{D}'(S^*M)$. Consequently, there is a non-negative Borel measure μ_h on S^*M such that

$$u_h(a_0) = \int_{S^*M} a_0 \, d\mu_h \,, \quad a_0 \in C^\infty(S^*M).$$

Proof. Since the map $a_0 \mapsto u_h(a_0)$ is a linear map from $C^\infty(S^*M)$ to $\mathbb{C}$, we would conclude that $u_h \in \mathcal{D}'(S^*M)$ if we could show that there is a constant C_h depending only on our fixed $h \in L^2(M)$ such that whenever $a_0 \in C^*\infty(S^*M)$ is real-valued, we had

$$|u_h(a_0)| \leq C_h \sup_{(x,\xi)\in S*M} |a_0(x,\xi)|. \tag{20.58}$$

To prove this we note that

$$a_0^\pm(x,\xi) := \sup|a_0| + a_0(x,\xi) \geq 0.$$

If $a^\pm \in C^\infty(T^*M \setminus 0)$ denotes the homogeneous of degree zero extension of $a_0^\pm$, then, by (20.57)

$$(h, a_F^\pm(x,D)h) \geq 0. \tag{20.59}$$

Let $\mathbf{1}_F(x,D)$ denote the ΨDO of order zero given by Theorem 20.43 when the symbol is identically one. Then

$$a_F^\pm(x,D) = \sup|a_0| \mathbf{1}_F(x,D) \pm a_F(x,D).$$

Therefore, by (20.59)

$$\sup|a_0|(h, \mathbf{1}_F h) \pm u_h(a_0) \geq 0,$$

and so

$$|u_h(a_0)| \leq (h, \mathbf{1}_F(x,D)h) \sup|a_0|. \tag{20.60}$$

Since zero-order ΨDOs are L^2 bounded, we obtain by Cauchy–Schwarz

$$|(h, \mathbf{1}_F(x,D)h)| \lesssim \|h\|_{L^2(M)}^2,$$

which means that (20.58) is indeed valid. Thus, $u_h \in \mathcal{D}'(S^*M)$.

Since (20.57) implies that u_h is non-negative, the last part of the assertion follows from Schwartz' theorem saying that non-negative distributions coincide with Borel measures. □

Note that if μ_h is the above Borel measure, associated to h, then, by (20.60) with $a_0 \equiv 1$, we have the following bound for its mass, namely

$$\mu_h(S^*M) = \int_{S^*M} d\mu_h \leq \|\mathbf{1}_F(x,D)\|_{L^2\to L^2} \|h\|_{L^2(M)}^2.$$

21. Introduction to ℓ^2 decoupling and some applications

In Section 8 we already saw that the square function conjecture (Conjecture 8.2)

$$\|f\|_{L^{2d/(d-1)}(\mathbb{R}^d)} \lesssim_\varepsilon R^\varepsilon \left\| \left(\sum_{\theta:R^{-1/2}-\text{slab}} |f_\theta|^2 \right)^{1/2} \right\|_{L^{2d/(d-1)}(\mathbb{R}^d)} \tag{21.1}$$

for all $f \in \hat{\mathcal{S}}(\mathbb{R}^d)$ with Fourier support in $\mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1})$, together with the Kakeya conjecture (Conjecture 15.1 in the form (8.2)) implies the restriction conjecture. Although we did not discuss this so far, an argument of Carbery [38] in fact shows that the hypothesized square function estimate (21.1) implies the Kakeya conjecture and, consequently, the restriction conjecture. [25]

In this section, we will therefore consider a weaker "analog" of (21.1) which is known as *ℓ^2-decoupling inequality*

$$\|f\|_{L^p(\mathbb{R}^d)} \lesssim_\varepsilon R^\varepsilon \left(\sum_{\theta:R^{-1/2}-\text{slab}} \|f_\theta\|^2_{L^p(\mathbb{R}^d)} \right)^{1/2} , \tag{21.2}$$

where the order of the mixed-norms on the right sides of (21.1) are now interchanged. The idea of this inequality is similar to the usual square function inequality, namely, it tries to separate or *decouple* the different frequency portions f_θ (contributing to $\|f\|_p$) from each other. This is done in an efficient as possible way to take the cancellations between the f_θ into account. In this regard however, (21.2) is clearly weaker than (21.1) by the triangle inequality for $2 \leq p (\leq 2d/(d-1))$, since

$$\|(\sum_\theta |f_\theta|^2)^{1/2}\|^2_{L^p} = \|\sum_\theta |f_\theta|^2\|_{L^{p/2}} \leq \sum_\theta \|f_\theta\|^2_{L^p} .$$

Moreover, we emphasize that (21.2) *does not* act as a substitute for (21.1) in the sense that it is not clear that it would imply the Kakeya or even the restriction conjecture. However, besides the fact that the right side of (21.2) is much easier to compute than the right side of (21.1) (as only size considerations will have to be made), decoupling theory does have a plethora of applications in PDE, additive combinatorics and number theory, see, e.g, the discussion in Carbery [38].

To simplify the upcoming notation, we make the following

Definition 21.1 (Decoupling norms)**.** For $1 \leq p \leq \infty$ and $f \in \hat{\mathcal{S}}(\mathbb{R}^d)$, we denote the *p-th decoupling norm* by

$$\|f\|_{L^{p,R^{-1}}(\mathbb{R}^d)} := \left(\sum_{\theta:R^{-1/2}-\text{slab}} \|f_\theta\|^2_{L^p(\mathbb{R}^d)} \right)^{1/2} .$$

For $\Omega \subseteq \mathbb{R}^d$ with finite Lebesgue measure, we analogously define the *local decoupling norms*

$$\|f\|_{L^{p,R^{-1}}(\Omega^d)} := \left(\sum_{\theta:R^{-1/2}-\text{slab}} \|f_\theta\|^2_{L^p(\Omega^d)} \right)^{1/2} .$$

[25] Attempting to prove the whole restriction conjecture from this point seems a quite optimistic strategy as (8.1) appears to be very powerful and in all likelihood considerably more difficult than the restriction conjecture.

and

$$\|f\|_{L^{p,R^{-1}}_{\mathrm{avg}}(\Omega^d)} := \left(\sum_{\theta:R^{-1/2}-\mathrm{slab}} \|f_\theta\|^2_{L^p_{\mathrm{avg}}(\Omega^d)} \right)^{1/2} .$$

where we recall $\|f\|_{L^p_{\mathrm{avg}}(\Omega)} = \|f\|_{L^p(\Omega,|\Omega|^{-1}dx)} = |\Omega|^{-1/p}\|f\|_{L^p(\Omega)}$.

In this notation, (21.2) takes the following form.

Theorem 21.2 (ℓ^2-decoupling)**.** *With the above notation,*

$$\|f\|_{L^p(\mathbb{R}^d)} \lesssim_\varepsilon R^{\varepsilon+\alpha(p)}\|f\|_{L^{p,R^{-1}}(\mathbb{R}^d)} \tag{21.3}$$

holds for all $f \in \mathcal{S}(\mathbb{R}^d)$ *with Fourier support in* $\mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1})$ *and* $2 \leq p \leq \infty$ *where*

$$\alpha(p) := \begin{cases} 0 & \text{if } 2 \leq p \leq 2(d+1)/(d-1) \\ (d-1)/4 - (d+1)/(2p) & \text{if } p > 2(d+1)/(d-1) \end{cases}$$

This theorem was already somewhat anticipated by Wolff [196] (in ℓ^p with p not necessarily 2) and proven for the first time by Garrigós–Seeger [91]. Bourgain [22] obtained the result for $2 \leq p \leq 2d/(d-1)$ and later, Bourgain and Demeter [29] (see also their study guide [30]) proved the inequality for the total "super-critical regime" $p \geq 2(d+1)/(d-1)$ (i.e., exponents above the Tomas–Stein restriction endpoint) [26]. Partial results in the super-critical regime were already obtained earlier by Demeter [62].

Albeit the exponent $2d/(d-1)$ plays a major role in the proof of the restriction conjecture, it turned out that this exponent is no longer optimal when considering the weaker decoupling inequalities; in fact, a more appropriate endpoint is the Tomas–Stein endpoint $2(d+1)/(d-1)$. For larger values of p, the obtained decoupling inequalities necessarily deteriorate when $R \to \infty$. (In fact, the polynomial behavior in R is optimal!) In applications it is often necessary to have the full power of Theorem 21.2 and, after discussing the preliminary estimate (21.2), we will detail how the complete range of estimates was proved later.

They key tool in the proof of Theorem 21.2 is multilinear restriction theory, which is well developed thanks to the work of Bennett–Carbery–Tao [7], see also Subsection 7.7. Before we discuss the proof in detail, let us have a brief look at some applications.

21.1. A first glimpse at applications.

21.1.1. *The discrete restriction phenomenon.* Recall the Tomas–Stein estimate for the paraboloid

$$\|\hat{f}\|_{L^2(\mathbb{P}^{d-1})} \lesssim \|f\|_{L^p(\mathbb{R}^d)}\,, \quad 1 \leq p \leq \frac{2(d+1)}{d+3}$$

which is, via localization theory, equivalent to

$$\|F\|_{L^p(B(R^{1/2}))} \lesssim R^{-1/4}\|\hat{F}\|_{L^2(\mathcal{N}_{R^{-1/2}}(\mathbb{P}^{d-1}))}\,, \quad p \geq \frac{2(d+1)}{d-1}$$

for any F with $\hat{F} \in C^\infty(\mathcal{N}_{R^{-1/2}}(\mathbb{P}^{d-1}))$ (see Lemma 6.3). Since F is localized to a ball of radius $R^{1/2}$ it is natural to expect that $\hat{F}$ is constant on the scale $R^{-1/2}$ and to approximate $\hat{F}$ by a weighted sum of indicator functions of balls of radius $R^{-1/2}$, i.e.,

$$\hat{F} \sim \sum_{\eta\in\Lambda} \hat{F}(\eta)\mathbf{1}_{B_\eta(R^{-1/2})}\,,$$

[26] The subcritical estimates follow from the $p = 2(d+1)/(d-1)$ case together with the trivial $p = 2$ inequality. The details of this argument will be discussed later.

where $\Lambda \subseteq \mathbb{P}^{d-1}$ is a maximal $R^{-1/2}$-separated subset (think of a lattice as a first approximation). Since we are only really interested in the values of $\hat{F}$ at the vertices of Λ, we can push this further and consider expressions of the form

$$\sum_{\eta\in\Lambda} a(\eta)\delta_\eta$$

where $a(\eta) \in \mathbb{C}$ are coefficients (weights) and δ_η is a Dirac δ mass concentrated at η. The inverse Fourier transform of such an expression therefore becomes a trigonometric polynomial, and so we see, heuristically at least, that the original Tomas–Stein estimate has the following discrete analog corresponding to an exponential sum estimate.

Corollary 21.3 (Discrete Tomas–Stein restriction theorem)**.** *For any maximal $\delta^{1/2} := R^{-1/2}$ separated set $\Lambda \subseteq \mathbb{P}^{d-1}$ and any $a : \Lambda \to \mathbb{C}$, the extension estimate*

$$\left\| \sum_{\eta\in\Lambda} a(\eta) \mathrm{e}^{2\pi i \langle \cdot, \eta\rangle} \right\|_{L^p_{\mathrm{avg}}(B(R^{1/2}))} \lesssim \delta^{\frac{d}{2p} - \frac{d-1}{4}} \|a\|_{\ell^2(\Lambda)}\,, \quad p \geq \frac{2(d+1)}{d-1}\,. \tag{21.4}$$

Remarks 21.4. (1) In fact, the discrete restriction estimate (21.4) is equivalent to the classical Tomas–Stein estimate, see, e.g., Demeter [63, Propositions 1.29 and 1.37] for the converse of what we proved here. Thus, the Tomas–Stein estimate measures the L^p-average of frequency-separated exponential sums at a spatial scale which is reciprocal to the separation of the frequencies. Observe also that (21.4) gives an improvement of $\delta^{d/(2p)}$ over the Cauchy–Schwarz inequality (which corresponds to the situation when no oscillation/cancellation is present).

(2) In the early 90s, Vega already proved a discrete analog of the Stein–Tomas–Strichartz restriction theorem. We recall

Theorem 21.5 (Vega [193, Theorem 3])**.** *Let $N \in \mathbb{N}$ and $m \in \mathbb{Z}^{d-1}$, $d \geq 2$. Then*

$$\left(\int_{|t|\leq N^{-1}} \int_{\mathbb{T}^{d-1}} \left| \sum_{i=1}^{d-1} \sum_{m_i=1}^{N} a_m \mathrm{e}^{it|m|^2} \mathrm{e}^{im\cdot x} \right|^p dx\, dt \right)^{1/p} \leq C_{p,N} \left(\sum_m |a_m|^2 \right)^{1/2}$$

holds where

$$C_{p,N} = \begin{cases} C_p N^{\frac{d-1}{2} - \frac{d+1}{p}} & \text{if } p > \frac{2(d+1)}{d-1}\,, \\ C & \text{if } p = \frac{2(d+1)}{d-1}\,, \\ C_p N^{-\frac{1}{2} + \frac{d+1}{2}\left(\frac{1}{2} - \frac{1}{p}\right)} & \text{if } 2 \leq p < \frac{2(d+1)}{d-1}\,, \end{cases}$$

and C_p are constants independent of N.

(3) This corollary and the ensuing Theorem 21.3 also hold when $\mathbb{P}^{d-1}$ is replaced by $\mathbb{S}^{d-1}$, but see also Bourgain–Demeter [29, Theorem 2.2].

Proof of Corollary 21.3. Without loss of generality, we assume that $B(R^{1/2})$ is centered at the origin. Let us now fix $\psi \in \widehat{C_c^\infty}(\mathbb{R}^d)$ with $\operatorname{supp}\psi \subseteq B_0(1)$ and $|\check{\psi}(x)| \gtrsim 1$ for $x \in B_0(1)$. As usual, let $\psi_{R^{-1/2}}(\xi) := R^{d/2}\psi(R^{1/2}\xi)$. Abbreviating

$$F := \sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i\langle\cdot,\eta\rangle}\,,$$

applying the localized Tomas-Stein estimate, and observing that the summands in

$$\widehat{F\check{\psi}_{R^{-1/2}}}(\xi) = \sum_{\eta\in\Lambda} a(\eta)\psi_{R^{-1/2}}(\xi - \eta)$$

have pairwise disjoint Fourier support (by the separation hypothesis on Λ and the definition of ψ) contained in $\mathcal{N}_{R^{-1/2}}(\mathbb{P}^{d-1})$, yields

$$\begin{aligned}\Big\|\sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i\langle\cdot,\eta\rangle}\Big\|_{L^p(B(R^{1/2}))} &\lesssim \|F\check{\psi}_{R^{-1/2}}\|_{L^p(B(R^{1/2}))} \lesssim R^{-1/4}\|\widehat{F\check{\psi}_{R^{-1/2}}}\|_{L^2(\mathcal{N}_{R^{-1/2}}(\mathbb{P}^{d-1}))}\\ &= R^{-1/4}\left(\sum_{\eta\in\Lambda}|a(\eta)|^2\int_{\widehat{\mathbb{R}^d}}|\psi_{R^{-1/2}}(\xi-\eta)|^2\,d\xi\right)^{1/2}\\ &= R^{-1/4+d/4}\left(\sum_{\eta\in\Lambda}|a(\eta)|^2\int_{\widehat{\mathbb{R}^d}}|\psi(\xi-\eta)|^2\,d\xi\right)^{1/2}\\ &\lesssim R^{(d-1)/4}\|a\|_{\ell^2(\Lambda)} = \delta^{-(d-1)/4}\|a\|_{\ell^2(\Lambda)}\,.\end{aligned}$$

(The scaling $\xi\mapsto R^{-1/2}$ from the second to the third line yields a factor of $R^{-d/2}$. Moreover, the support of $\psi_{R^{-1/2}}$, i.e., roughly $\mathcal{N}_{R^{-1/2}}(\mathbb{P}^{d-1})$, is transformed into $\mathcal{N}_1(\mathbb{P}^{d-1})$.) The claim follows now from the definition of the L^p_{avg} norm which yields the "missing" $\delta^{d/(2p)}$ factor. □

Now, Bourgain and Demeter made the fundamental observation that, as soon as one averages in physical space over much larger balls, one obtains improvements over the classical Tomas–Stein inequality because of additional cancellations (through oscillations). These cancellations are a consequence of the ℓ^2-decoupling as we will see now.

Theorem 21.6 (Discrete restriction phenomenon)**.** *Let $\Lambda\subseteq\mathbb{P}^{d-1}$ be a maximal $\delta^{1/2}$-separated subset, $a:\Lambda\to\mathbb{C}$, and $R\geq\delta^{-1}$. Then, for all $\varepsilon>0$, we have the extension estimate*

$$\Big\|\sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i\langle\cdot,\eta\rangle}\Big\|_{L^p_{\mathrm{avg}}(B(R))} \lesssim_\varepsilon \delta^{\frac{d}{2p}-(d-1)/4+1/(2p)-\varepsilon}\|a\|_{\ell^2(\Lambda)}\,,\quad p\geq\frac{2(d+1)}{d-1}\,. \tag{21.5}$$

Remark 21.7. Observe two things.

(1) $R\geq\delta^{-1}$ is now rather variable. But more importantly,
(2) we are now averaging over balls with the much larger radius R (instead of $R^{1/2}$). This averaging over larger balls is precisely the source of the $\delta^{1/2p}$-improvement over the classical Tomas–Stein inequality.

Proof. Let us prepare the proof with some preliminaries. Fix an R-ball $B_R = B_{x_0}(R)$ and let $\psi\in\widehat{C_c^\infty}(\mathbb{R}^d)$ be as in the previous proof with Fourier support contained in $B_0(1)$ and $\psi_R(\xi)=R^d\psi(R\xi)$. Let furthermore $g:\mathbb{P}^{d-1}\to\mathbb{R}$ be a nice function and observe that $(gd\sigma)*\psi_R$ has Fourier support contained in the R^{-1}-neighborhood $\mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1})\subseteq\mathcal{N}_\delta(\mathbb{P}^{d-1})$. Clearly, the left side of the classical, localized, Thomas–Stein estimate can be bounded by

$$\int_{B_R}|(gd\sigma)^\vee(x)|^p\,dx \lesssim \int_{B_R}|((gd\sigma)*\psi_R)^\vee(x)|^p\,dx = \int_{B_R}|(gd\sigma)^\vee\cdot\check{\psi}_R(x)|^p\,dx\,.$$

Now, applying the ℓ^2-decoupling inequality (21.3) to $f:=((gd\sigma)*\psi_R)^\vee$, we obtain (with the previous estimate)

$$\left(\frac{1}{|B_R|}\int_{B_R}|(gd\sigma)^\vee(x)|^p\,dx\right)^{1/p} \lesssim_\varepsilon R^{\varepsilon-d/p}\cdot\delta^{\frac{d+1}{2p}-(d-1)/4}\left(\sum_{\theta:\delta^{1/2}-cap}\|(g_\theta d\sigma)^\vee\cdot\check{\psi}_R\|_p^2\right)^{1/2} \tag{21.6}$$

where $g_\theta := g\mathbf{1}_\theta$ is the restriction of g onto the cap $\theta\cap\mathbb{P}^{d-1}$. (Recall that for $p \geq 2(d+1)/(d-1)$, we had $\alpha(p) = (d-1)/4 - (d+1)/(2p)$ in the decoupling inequality, which is just the negative exponent of δ on the right side of this formula.)

For a given $\varepsilon > 0$ and $\eta \in \mathbb{P}^{d-1}$, let $P(\eta,\varepsilon) := \mathbb{P}^{d-1} \cap B_\eta(\varepsilon)$ be an arbitrary ε-cap, centered at η of the paraboloid and consider the function

$$g^\varepsilon := \sum_{\eta\in\Lambda} a(\eta)\frac{1}{\sigma(P(\eta,\varepsilon))}\mathbf{1}_{P(\eta,\varepsilon)}\,,$$

where we recall that $\sigma(P(\eta,\varepsilon))$ was the euclidean surface measure of the set $P(\eta,\varepsilon)$ on $\mathbb{P}^{d-1}$. Now, observe first that $(g_\varepsilon d\sigma)^\vee(x)$ converges pointwise to the function on the left side of our assertion (e.g., by Lebesgue's differentiation theorem), i.e.,

$$\lim_{\varepsilon\to0}(g^\varepsilon d\sigma)^\vee(x) = \lim_{\varepsilon\to0}\sum_{\eta\in\Lambda} a(\eta)\frac{1}{\sigma(P(\eta,\varepsilon))}\int_{P(\eta,\varepsilon)} \mathrm{e}^{2\pi i x\cdot\xi}\, d\sigma(\xi) = \sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i x\cdot\eta}\,.$$

Thus, by Fatou's lemma, i.e.,

$$\|\sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i\langle\cdot,\eta\rangle}\|_p \leq \liminf_{\varepsilon\to0}\|(g^\varepsilon d\sigma)^\vee\|_p$$

it suffices to estimate further the right side of (21.6) with g_θ replaced by g_θ^ε. First, for $\varepsilon \ll \delta^{1/2}$ (think of $\varepsilon = R^{-1}$), we have the pointwise estimate

$$\begin{aligned}|(g_\theta^\varepsilon d\sigma)^\vee(x)| &= \left|\sum_{\eta\in\Lambda} a(\eta)\frac{1}{\sigma(P(\eta,\varepsilon))}\int_{P(\eta,\varepsilon)\cap\theta} \mathrm{e}^{2\pi i x\cdot\xi}\, d\sigma(\xi)\right| \leq \sum_{\eta\in\Lambda\,,P(\eta,\varepsilon)\cap\theta\neq\emptyset} |a(\eta)| \\ &\lesssim \left(\sum_{\eta\in\Lambda\,,P(\eta,\varepsilon)\cap\theta\neq\emptyset} |a(\eta)|^2\right)^{1/2}\end{aligned}$$

where we used Cauchy–Schwarz together with the fact that

$$\#\{\eta\in\Lambda : P(\eta,\varepsilon)\cap\theta\neq\emptyset\} = \mathcal{O}(1)\,,$$

because Λ is a maximal $\delta^{1/2}$-separated set, θ is an $\delta^{1/2}$-cap, and $P(\eta,\varepsilon)$ can intersect with at most one such slab as $\varepsilon \ll \delta$. Plugging this estimate in the L^p norm of the right side of (21.6) yields

$$\begin{aligned}\left(\sum_{\theta:\delta^{1/2}-cap} \|(g_\theta^\varepsilon d\sigma)^\vee\check{\psi}_R\|^2_{L^p(\mathbb{R}^d)}\right)^{1/2} &\lesssim \left(\sum_{\theta:\delta^{1/2}-cap}\ \sum_{\eta\in\Lambda,\, P(\eta,\varepsilon)\cap\theta\neq\emptyset} |a(\eta)|^2\|\check{\psi}_R\|^2_{L^p(\mathbb{R}^d)}\right)^{1/2} \\ &\lesssim R^{d/p}\left(\sum_{\theta:\delta^{1/2}-cap}\ \sum_{\eta\in\Lambda,\, P(\eta,\varepsilon)\cap\theta\neq\emptyset} |a(\eta)|^2\right)^{1/2} \\ &\lesssim R^{d/p}\|a\|_{\ell^2(\Lambda)}\end{aligned}$$

where we used in the final inequality that the cardinality of the θ-sum is of order $\mathcal{O}(1)$ for fixed η because $\varepsilon \ll \delta^{1/2}$ and θ is a $\delta^{1/2}$-cap. This concludes the proof of the theorem. □

21.1.2. *Strichartz estimates for the Schrödinger equation on the torus.* We follow the notes of Hickman and Vitturi [107, p. 22, Lecture 2, Section 2.2].

As we have already seen in Subsection 13.1, restriction estimates immediately imply estimates for solutions of dispersive PDE posed in $\mathbb{R}^d$. It is natural to generalize these ideas to PDEs posed on finite domains with certain boundary conditions. Here, we focus on the unit cube with periodic boundary conditions, more precisely on the Schrödinger equation on the torus $\mathbb{T}^d = \mathbb{R}^d \setminus \mathbb{Z}^d$. In the early 90's Bourgain [18] (but see also [21] for irrational tori and a "survey") found that the solution of the Schrödinger equation includes waves which travel with different directions around the torus. As one may imagine, it is very challenging to estimate how these different waves interfere with each other and to find estimates on them. At that time Bourgain could prove sharp estimates only in $d = 2, 3$. Surprisingly, the analysis required many tools from number theory. For instance, it uses unique factorization of integers in order to estimate the number of solutions of some diophantine equations. For higher dimensions, the problem seemed out of reach and it was supposed that the solution required both Fourier analysis and number theory. Bourgain and Demeter found that decoupling inequalities were the crucial tool to obtain dispersive estimates in higher dimensions.

Clearly, dispersive estimates for the solution

$$u(x,t) := \mathrm{e}^{-i(2\pi)^{-1}t\Delta}\varphi(x) := \sum_{\xi\in\mathbb{Z}^d} \hat{\varphi}(\xi)\mathrm{e}^{2\pi i(x\cdot\xi + t\xi^2)}$$

of the Schrödinger equation on $\mathbb{T}^d \times \mathbb{R}$ are obtained using the previously discussed discrete restriction estimates. Now, due to the above discussion, i.e., the fact that a general solution consists of many waves traveling in different directions, we can certainly *not* expect the original Strichartz estimates for the equation on $\mathbb{R}^d$ to hold. In fact, Bourgain [18] proved the failure of Strichartz estimates on $\mathbb{T}^1$. (Observe that the exponent $q = 6$ really is the Strichartz exponent in $d = 1$, see Theorem 13.1.)

Theorem 21.8 (Failure of Strichartz on $\mathbb{T}^1 \times [0,1]$)**.** *For every $N \in \mathbb{N}$ there exists a smooth function φ_N on $\mathbb{T}$ with* supp $\hat{\varphi}_N \subseteq [-N, N]$ *such that*

$$\|\mathrm{e}^{-i(2\pi)^{-1}t\Delta}\varphi_N\|_{L^6(\mathbb{T}\times[0,1])} \gtrsim (\log N)^{1/6}\|\varphi_N\|_{L^2(\mathbb{T})} \,. \tag{21.7}$$

In particular, we could take $\hat{\varphi}_N(\xi) = \mathbf{1}_{\{0,1,\dots,N\}}(\xi)$ (i.e., φ_N is a trigonometric polynomial) so that we are in the situation of discrete restriction phenomena, i.e.,

$$\mathrm{e}^{-i(2\pi)^{-1}t\Delta}\varphi_N(x) = \sum_{n=0}^{N} \mathrm{e}^{2\pi i(xn+tn^2)} \,.$$

This solution is known as a *Weyl sum* (or *Gauss sum*, see also Bourgain's counterexample [23] for the a.e. convergence of solutions to the Schrödinger equation) and it is of considerable interest in number theory. In fact, the lower bound in (21.7) can be obtained by appealing to number-theoretic techniques (such as the Hardy–Littlewood–Ramanujan circle method).

Now, the question is whether one can nevertheless establish Strichartz estimates with a sharp dependence on the size of the frequency support of the initial data. For instance, in view of the above counterexample, we may pose the

Question: "Can one prove an $L^2_x(\mathbb{T}) \to L^6_{x,t}(\mathbb{T}\times[0,1])$ Strichartz estimate for initial data φ_N with supp $\varphi_N \subseteq [-N, N]$ but with a *sub-polynomial dependence on* N?"

Fortunately, with the help of the discrete restriction estimates proved above, we have

Theorem 21.9 (Strichartz on $\mathbb{T}^d$ [29]). *Let $\varphi \in L^2(\mathbb{T}^d)$ with* supp $\hat{\varphi} \subseteq [-N,N]^d$*. Then for any time interval $I \subseteq \mathbb{R}$ with $|I| \gtrsim 1$, we have for any $\varepsilon > 0$,*

$$\|\mathrm{e}^{-i(2\pi)^{-1}t\Delta}\varphi\|_{L^p(\mathbb{T}^d\times I)} \lesssim_\varepsilon N^{d/2-(d+2)/p+\varepsilon}|I|^{1/p}\|\varphi\|_{L^2(\mathbb{T}^d)}\,, \quad p \geq \frac{2(d+2)}{d}\,. \tag{21.8}$$

Up to the subpolynomial loss, Theorem 21.9 is sharp. As we have outlined in the beginning of this subsubsection, the earlier partial results in higher dimensions were crucially based on number theoretic arguments which will not be able in the following argument. In particular, it seems that the current techniques are more robust; in particular, one can apply the following argument also to the analogous problem posed on "irrational tori", see Bourgain–Demeter [29].

Proof of Theorem 21.9. To ease the notation and make the connection with Theorem 21.6 clear, we set $n = d+1$. For $\xi' \in \mathbb{Z}^{n-1}$ with $|\xi'|_\infty \leq N$, let $\eta' := N^{-1}\xi'$ and $\eta_n := |\eta'|^2$ so that the collection Λ of all $\eta = (\eta',\eta_n)$ becomes a (maximal) N^{-1}-separated subset of $\mathbb{P}^{n-1}$. Defining $a(\eta) := \hat{\varphi}(N\eta')$ and scaling ($x \mapsto x/N$ and $t \mapsto t/N^2$), we obtain

$$\left(\int_{\mathbb{T}^{n-1}\times I} |\mathrm{e}^{-i(2\pi)^{-1}t\Delta}\varphi(x)|^p\,dx\,dt\right)^{1/p} = N^{-(n+1)/p}\left(\int_D \left|\sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i y\cdot\eta}\right|^p dy\right)^{1/p} \tag{21.9}$$

where the domain of integration D is given by

$$D := \{y \in \mathbb{R}^n : \ |y_j| \leq N/2 \text{ for } 1 \leq j \leq n-1 \text{ and } y_n \in N^2 I\}\,,$$

and we identified $\mathbb{T}$ with $[-1/2,1/2]$ for convenience.

We will now estimate the right side of (21.9) from above by a localized L^p norm on some ball of radius $\sim N^2$ to apply (21.5). Since $\eta' \in N^{-1}\mathbb{Z}^{n-1}$ for each $\eta \in \Lambda$, the above integrand is periodic with period N in the variables y'. Now, let $R := N^2|I| \gtrsim N^2 =: \delta^{-1}$ and $B_R := B_{\frac{N^2|I|}{2}e_n}(R)$. Note that B_R can be covered by $\mathcal{O}((|I|N)^{n-1})$ sets of the form $D + N(k',0)$ where $k' \in \mathbb{Z}^{n-1}$. These observations allow us to estimate (21.9) from above by

$$|I|^{1/p}\left(\frac{1}{|B_R|}\int_{B_R}\left|\sum_{\eta\in\Lambda} a(\eta)\mathrm{e}^{2\pi i y\cdot\eta}\right|^p dy\right)^{1/p} \lesssim_\varepsilon |I|^{1/p}N^{(n-1)/2-(n+1)/p+\varepsilon}\|a\|_{\ell^2(\Lambda)}$$

where we used the discrete restriction phenomenon (21.5) with $\delta = N^{-2}$. Since $\|a\|_{\ell^2(\Lambda)} = \|\hat{\varphi}\|_{L^2(\mathbb{Z}^{n-1})}$ by the definition of $a(\eta)$ and Plancherel, the theorem is proved. ☐

21.2. **Some preliminary observations.**

Definition 21.10. Let $p \in [1,\infty]$ and $\mathcal{U} = \{U_1,...,U_n\}$ be a finite collection of non-empty subsets of $\mathbb{R}^d$ for some $n \geq 1$. (We permit repetitions, so $\mathcal{U}$ is in fact $\mathcal{U}$ may rather be a multi-set than a set.) We define the decoupling constant $\mathrm{Dec}_p(\mathcal{U})$ to be the smallest constant for which there is an inequality

$$\|\sum_j f_j\|_{L^p(\mathbb{R}^d)} \leq \mathrm{Dec}_p(\mathcal{U})(\sum_j \|f_j\|^2_{L^p(\mathbb{R}^d)})^{1/2} \tag{21.10}$$

whenever $f \in \mathcal{S}(\mathbb{R}^d)$ has Fourier support in U_j.

Remarks 21.11. (1) We have the trivial bounds

$$1 \leq \mathrm{Dec}_p(\mathcal{U}) \leq n^{1/2}\,. \tag{21.11}$$

The upper bound follows from applying the triangle inequality and then Cauchy–Schwarz whereas the lower bound comes from taking just one f_j to be non-zero. Clearly, it would be very desirable to show $\mathrm{Dec}_p(\mathcal{U}) = \mathcal{O}_{p,d}(1)$, uniformly in n. However, the best, one can do at the moment is (because all so-far known proofs use an induction of scales argument) a subpolynomial loss, i.e., for any $\varepsilon > 0$, one has $\mathrm{Dec}_p(\mathcal{U}) \lesssim_\varepsilon n^\varepsilon$.

(2) In Proposition 8.4 we observed that the reverse square function estimate holds in L^2 and L^4 when we assume that the U_j (respectively the set-sums $U_i + U_j$) overlap only finitely. Thus, by the triangle inequality, we obtain in these cases that $\mathrm{Dec}_2(\mathcal{U}) \leq A_2^{1/2}$ and $\mathrm{Dec}_4(\mathcal{U}) \leq A_4^{1/4}$ where A_2, A_4 are defined in Proposition 8.4.

(3) In the literature, the U_j are often assumed to be pairwise disjoint. However, here it is convenient to allow them to be finitely overlapping to circumvent some minor technicalities.

Proposition 21.12 (Elementary properties of decoupling constants)**.** *Let $1 \leq p \leq \infty$ and $d \geq 1$. Then, the decoupling constant has the following properties.*

1. *(Monotonicity) We have $\mathrm{Dec}_p(\mathcal{U}) \leq \mathrm{Dec}_p(\mathcal{U}')$ whenever $\mathcal{U}' = \{U_j'\}_{j=1}^n$ is a collection whose elements contain U_j, i.e., $U_j \subseteq U_j'$ for all $j = 1, ..., n$.*
2. *(Triangle inequality) We have*
$$\mathrm{Dec}_p(\mathcal{U}), \mathrm{Dec}_p(\mathcal{U}') \leq \mathrm{Dec}_p(\mathcal{U} \cup \mathcal{U}') \leq (\mathrm{Dec}_p(\mathcal{U})^2 + \mathrm{Dec}_p(\mathcal{U}')^2)^{1/2}$$
for all non-empty collections $\mathcal{U}, \mathcal{U}'$ of open, non-empty subsets of $\mathbb{R}^d$.
3. *(Affine invariance) Let $U_1, ..., U_n$ be non-empty, open subsets of $\mathbb{R}^d$ and $L : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible affine transformation. Then, we have $\mathrm{Dec}_p(LU_1, ..., LU_n) = \mathrm{Dec}_p(U_1, ..., U_n)$.*
4. *(Interpolation) Let $1/p = (1-\theta)/p_0 + \theta/p_1$ for $1 \leq p_0 \leq p \leq p_1 \leq \infty$ and $0 \leq \theta \leq 1$. Suppose that we have for $\mathcal{U} = \{U_1, ..., U_n\}$ (with $U_j \subseteq \mathbb{R}^d$ non-empty, open) the projection bounds*
$$\|P_{U_j} f\|_{L^{p_i}(\mathbb{R}^d)} \lesssim_{p_i,d} \|f\|_{L^{p_i}(\mathbb{R}^d)}, \quad i = 0, 1, \; j = 1, ..., n, \; f \in \mathcal{S}(\mathbb{R}^d),$$
where the Fourier multiplier P_{U_j} is defined by
$$\widehat{P_{U_j} f}(\xi) := \mathbf{1}_{U_j}(\xi) \hat{f}(\xi).$$
Then we have
$$\mathrm{Dec}_p(\mathcal{U}) \lesssim_{p_0,p_1,d,\theta} \mathrm{Dec}_{p_0}(\mathcal{U})^{1-\theta} \mathrm{Dec}_{p_1}(\mathcal{U})^\theta.$$
5. *(Multiplicativity) Suppose that $\mathcal{U} = \{U_1, ..., U_n\}$ is a collection of non-empty open subsets of $\mathbb{R}^d$ where each U_j is partinitioned (up to null-sets) into $U_j = \bigcup_{\ell=1}^{m_j} U_{j,\ell}$ for some disjoint non-empty open subsets of $\mathbb{R}^d$. If $p \geq 2$, then*
$$\mathrm{Dec}_p(\{U_{j,\ell} : \; j = 1, ..., n, \ell = 1, ..., m_j\}) \leq \mathrm{Dec}_p(\mathcal{U}) \times \sup_{j \in \{1,...,n\}} \mathrm{Dec}_p(\{U_{j,1}, ..., U_{j,m_j}\})$$
6. *(Adding trivial dimensions) Suppose that $\{U_1, ..., U_n\}$ is a collection of non-empty open subsets of $\mathbb{R}^d$ and $p \geq 2$. Then, for any $d' \geq 1$, we have*
$$\mathrm{Dec}_p(U_1, ..., U_n) = \mathrm{Dec}_p(U_1 \times \mathbb{R}^{d'}, ..., U_n \times \mathbb{R}^{d'})$$
where the right side is the decoupling constant in $\mathbb{R}^d \times \mathbb{R}^{d'} = \mathbb{R}^{d+d'}$.

Proof. □

The following observation shows that there can be no ℓ^2 decoupling for an infinite partition in Fourier space, i.e., when $n \to \infty$.

Proposition 21.13. *Let $\mathcal{U} = \{U_1, .., U_n\}$ be a collection of non-empty open subsets in $\mathbb{R}^d$. Then, we have $\mathrm{Dec}_p(\mathcal{U}) \gtrsim n^{\frac{1}{p} - \frac{1}{2}}$. Equivalently, there exist smooth f_j with $\operatorname{supp} f_j$ contained in compact subsets of U_j such that*
$$\|\sum_{j=1}^n f_j\|_{L^p(\mathbb{R}^d)} \gtrsim n^{\frac{1}{p} - \frac{1}{2}} (\sum_{j=1}^n \|f_j\|_{L^p(\mathbb{R}^d)}^2)^{1/2}$$

for any $1 \leq p \leq 2$ *and the implicit constant does not depend on* $\mathcal{U}$ *or* n.

Proof. Set $\operatorname{supp} \hat{f}_j \subseteq B_{\eta_j}(\delta)$ for some $\eta_j \in U_j$ and $0 < \delta \ll 1$ and L^p-normalize the f_j. Next, we modulate the $\hat{f}_j$ such that the f_j are concentrated on balls $B_{x_j}(\delta^{-1})$ and decay rapidly away from these balls. That is, the $\hat{f}_j$ are of the form $\hat{f}_j(\xi) = \psi(\delta^{-1}(\xi - \eta_j))\mathrm{e}^{2\pi i x_j \cdot \xi}$ for some $\psi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} \psi \subseteq B_0(1)$. Moreover, we modulate the f_j such that $|x_j - x_i| \sim \delta^{-1}$ for any $i \neq j$. Therefore, we can bound

$$\|\sum_{j=1}^n f_j\|_p \gtrsim n^{1/p}$$

But since $(\sum_{j=1}^n \|f_j\|_{L^p(\mathbb{R}^d)}^2)^{1/2} \lesssim n^{1/2}$, this establishes the claim. □

Instead of modulating the f_j, we could have also randomized them in the spirit of Subsection 15.1.

Note that the the reverse triangle inequality in $L^{p/2}$ for $p < 2$ would have merely lead us to

$$(\mathbb{E}\|\sum_j \varepsilon_j f_j\|_p^p)^{2/p} \sim \|(\sum |f_j|^2)^{1/2}\|_p^2 = \|\sum |f_j|^2\|_{p/2} \geq \sum_j \||f_j|^2\|_{p/2} = \sum_j \|f_j\|_p^2 .$$

Remark 21.14. The above proof sheds also some light on why the Hausdorff–Young inequality $\|\hat{f}\|_q \lesssim \|f\|_p$ fails when $p > 2$, even when $q = p'$ (which is easily seen to be necessary by "dimensional analysis"). The idea is to have f "spread out" in physical space to keep the L^p the norm low. However, we would also like to spread out $\hat{f}$ in Fourier space to prevent the $L^{p'}$ norm from dropping too much. To this end, let

$$f(x) = \sum_{j=1}^n \varepsilon_j \varphi(x - x_j)$$

for random signs $\varepsilon_1, ..., \varepsilon_n$ and a non-zero bump function $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} \varphi \subseteq B_0(1)$. Here, we merely need that the centers x_j are sufficiently separated; $|x_i - x_j| \geq 2$ would do, for example. Since the summands are disjointly supported, we have on the one hand

$$\|f\|_p \sim n^{1/p} .$$

Thus, if Hausdorff–Young were true for $p > 2$, we would have the (probabilistic) bound $\|\hat{f}\|_{p'} \lesssim n^{1/p}$. But on the other hand, the Fourier transform is given by

$$\hat{f}(\xi) = \sum_{j=1}^n \varepsilon_j \mathrm{e}^{2\pi i x_j \cdot \xi} \hat{\varphi}(\xi)$$

and so by Khintchine's inequality $\mathbb{E}\|\hat{f}\|_{p'}^{p'} \sim \|(\sum_j |\mathrm{e}^{2\pi i \langle x_j, \cdot \rangle} \hat{\varphi}|^2)^{1/2}\|_{p'}^{p'}$, we have

$$\|\hat{f}\|_{p'} \sim \|(\sum_{j=1}^n |\hat{\varphi}|^2)^{1/2}\|_{p'} \sim n^{1/2}$$

which clearly contradicts $\|\hat{f}\|_{p'} \lesssim n^{1/p}$ unless $p \leq 2$. The point of the randomization argument is that it allows us to get rid of the phases $\mathrm{e}^{2\pi i x_j \cdot \xi}$ in Fourier space which could lead to substantial cancellations, thereby suppressing the L^q norm of $\hat{f}$.

Hence, we focus on ℓ^2-decoupling for $p \geq 2$ in what follows. We already saw that for $p = 2$, we obtained decoupling when the sets overlap only finitely. For larger p this constraint is insufficient as the next observation reveals. In particular, it tells us that we should require that the U_j are somewhat curved (in analogy to the restriction phenomenon).

Proposition 21.15. *If* $\mathcal{U} = \{(j, j+1) : \ 0 \leq j < n\}$*, and* $p \in [2,\infty]$*, then*

$$\mathrm{Dec}_p(\mathcal{U}) \sim n^{\frac{1}{p}-\frac{1}{2}}$$

Proof. □

21.3. **Uncertainty principles related to ℓ^2-decoupling.** Weighted estimates will be a common feature of our future analysis which motivates the following

Definition 21.16 (Smooth localization)**.** We denote by $w_{B_c(R)} \geq 0$ rapidly decaying weights concentrated on a ball $B_c(R)$, i.e., $w_{B_c(R)}$ satisfies $w_{B_c(R)}(x) \sim 1$ for $x \in B_c(R)$ and

$$w_{B_c(R)}(x) \lesssim \left(1 + \frac{|x-c|}{R}\right)^{-N} \quad \text{for some large } N = \mathcal{O}(1)\,.$$

The precise choice of $w_{B_c(R)}$ may vary from line to line or, indeed, within a single line. For various technical reasons it is preferable to work with this fairly general class of weights rather than with Schwartz functions. Let us also introduce the corresponding weighted norms.

Definition 21.17 (Smoothly localized norm)**.** For $p \in [1,\infty]$, let

$$\begin{aligned} &\|\cdot\|_{L^p(w_{B_c(R)})} \quad \text{and} \\ &\|\cdot\|_{L^p_{\mathrm{avg}}(w_{B_c(R)})} = \|\cdot\|_{L^p(|B_c(R)|^{-1}w_{B_c(R)})} = |B_c(R)|^{-1/p}\|\cdot\|_{L^p(w_{B_c(R)})} \end{aligned} \tag{21.12}$$

denote the L^p norms defined with respect to the measures $w_{B_c(R)}(x)\,dx$, respectively $|B_c(R)|^{-1}w_{B_c(R)}(x)\,dx$.

Let us state and prove the following local Bernstein inequality (cf. Proposition D.5) and orthogonality principles that will be invoked frequently later on.

Lemma 21.18 (Local Bernstein inequality / Reverse Hölder inequality)**.** *Let* $r \geq R \geq 0$*. If* f *satisfies* supp $\hat{f} \subseteq B_c(1/R)$*, then*

$$\|f\|_{L^q_{\mathrm{avg}}(B_c(r))} \lesssim (rR)^{d(1/p-1/q)}\|f\|_{L^p_{\mathrm{avg}}(w_{B_c(r)})}$$

holds for all $1 \leq p \leq q \leq \infty$*.*

Proof. We follow Hickman–Vitturi [107], but see also Demeter [63, Lemma 9.19]. For any such f we have the global Bernstein inequality (Proposition D.5)

$$\|f\|_{L^q(\mathbb{R}^d)} \lesssim R^{d(1/p-1/q)}\|f\|_{L^p(\mathbb{R}^d)}\,.$$

The local version follows by replacing f by $f\psi_{B_{c'}(r)}$ where $\psi_{B_{c'}(r)}$ is a modulated Schwartz function adapted to $B_{c'}(r)$ such that supp $\hat{\psi}_{B_{c'}(r)} \subseteq B_c(1/r) \subseteq B_c(1/R)$ and it holds that supp $\hat{f} * \hat{\psi}_{B_{c'}(r)} \subseteq B_c(2/R)$. □

Proposition 21.19 (Local orthogonality)**.** *For* $r \geq R^{1/2}$ *we have*

(1) $\|f\|_{L^2_{\mathrm{avg}}(B(r))} \lesssim \|f\|_{L^{2,R^{-1}}_{\mathrm{avg}}(w_{B(r)})}$ *and*
(2) $\|f\|_{L^2_{\mathrm{avg}}(w_{B(r)})} \lesssim \|f\|_{L^{2,R^{-1}}_{\mathrm{avg}}(w_{B(r)})}$,

whenever supp $\hat{f} \subseteq \mathcal{N}_{R^{-1}}(\mathbb{P}^{d-1})$.

This means, we can both control smoothly and non-smoothly localized L^2-averages by smoothly weighted decoupling norms (recall Definition 21.1).

Proof. (1) Let $\psi_{2r} \in \mathcal{S}(\mathbb{R}^d)$ such that $\psi_{2r}(x) \gtrsim 1$ for $x \in B(2r)$ and supp $\hat{\psi}_{2r} \subseteq B_0(1/(2r))$. Therefore,

$$\|f\|_{L^2_{\mathrm{avg}}(B(r))} \lesssim r^{-d/2}\|f\psi_{2r}\|_{L^2(\mathbb{R}^d)} = r^{-d/2}\|\sum_{\theta:R^{-1/2}-\mathrm{slab}} \widehat{f}_\theta * \hat{\psi}_{2r}\|_{L^2(\mathbb{R}^d)}\,.$$

This is already almost what we want. Now note that each $\widehat{f_\theta} * \hat{\psi}_{2r}$ is supported in $\mathcal{N}_{R^{-1/2}}(\theta)$. Moreover, since $r^{-1} \le R^{-1/2}$, we have that $\mathrm{supp}(\widehat{f_\theta} * \hat{\psi}_{2r})$ is contained in the union of only $\mathcal{O}(1)$ many $R^{-1/2}$-slabs. Thus, the $\mathrm{supp}(\widehat{f_\theta} * \hat{\psi}_{2r})$ overlap only finitely and therefore, we have

$$\|f\|_{L^2_{\mathrm{avg}}(B(r))} \lesssim \left(r^{-d} \sum_{\theta: R^{-1/2}-\mathrm{slab}} \|\widehat{f_\theta} * \hat{\psi}_{2r}\|_{L^2(\mathbb{R}^d)} \right)^{1/2} .$$

Using Plancherel and taking $w_{B(r)} := |\psi_{2r}|^2$ yields the desired estimate.

(2) We reduce to the first case by observing that

$$\|f\|^2_{L^2_{\mathrm{avg}}(w_{B(r)})} \lesssim \sum_{k\in\mathbb{Z}^d} (1+|k|)^{-N} \|f\|^2_{L^2_{\mathrm{avg}}(B(r)+kr)}$$

due to the rapid decay of $w_{B(r)}$. (Here $N = \mathcal{O}_d(1)$ is a large integer.) This allows us to apply part (1) of the proposition to each of the $\|f\|_{L^2_{\mathrm{avg}}(B(r)+kr)}$ to deduce

$$\|f\|^2_{L^2_{\mathrm{avg}}(w_{B(r)})} \lesssim \sum_{k\in\mathbb{Z}^d} (1+|k|)^{-N} \|f\|^2_{L^{2,R^{-1}}_{\mathrm{avg}}(w_{B(r)+kr})} .$$

Now, the right side is given by

$$\sum_{\theta: R^{-1/2}-\mathrm{slab}} r^{-d} \int_{\mathbb{R}^d} |f_\theta(x)|^2 \left(\sum_{k\in\mathbb{Z}^d} (1+|k|)^{-N} w_{B(r)+kr}(x) \right) dx .$$

But since the expression in parentheses is just another weight adapted to B_r, the right side equals

$$\sum_{\theta: R^{-1/2}-\mathrm{slab}} r^{-d} \int_{\mathbb{R}^d} |f_\theta(x)|^2 w_{B(r)}(x)\, dx = \|f\|^2_{L^{2,R^{-1}}_{\mathrm{avg}}(w_{B(r)})}$$

and we are done.

□

22. Summary

`[Proof of the restriction conjecture relies partly on understanding oscillatory integrals and on set theoretic problems, e.g., of Kakeya's type. See Tao [183, p. 298ff]]`

There are three "classic" (i.e., outdated) approaches to prove restriction estimates.

(1) Compute $(d\sigma)^\vee(x)$, perform a dyadic partition of unity of the kernel, and use interpolation to bound $\|f * (d\sigma)^\vee\|_{p'} \lesssim \|f\|_p$. This is the classic Tomas–Stein approach.

(2) Follow Strichartz' approach and compute the kernel of $(Q(-i\nabla) - z)^{-\zeta}$ where Q is the (quadratic?) form associated to S (e.g., $Q(\xi) = \xi^2$ or $Q(\xi) = -\xi_1^2 - \dots - \xi_j^2 + \xi_{j+1}^2 + \dots \xi_d^2$ for wave- (or Klein–Gordon)-like problems) for $\mathrm{Re}(\zeta) \ge 1$ (often $\mathrm{Re}\zeta \in [d/2, (d+1)/2]$).

(3) Go through the theory of inhomogeneous oscillatory integrals (see Theorem 4.6) where the Carleson–Sjölin conditions may not be met (Stein's and Bourgain's approach) and obtain the dual restriction (i.e., the extension) estimate as a corollary.

Appendix A. Selection of omitted proofs

A.1. **The ε-removal lemma.** We review the proof of Theorem 6.5 which is due to Tao [181, Theorem 1.2].

Theorem A.1. *Assume* $|(d\sigma)^\vee(x)| \lesssim (1+|x|)^{-\rho}$ *for some* $\rho > 0$*. If* $R_S(p \to p; \alpha)$ *holds for some* $p < 2$ *and* $0 < \alpha \ll 1$*, then one has* $R_S(q \to q)$ *whenever*

$$\frac{1}{q} > \frac{1}{p} + \frac{A_\rho}{\log(1/\alpha)}\,.$$

The first step is to bootstrap the localized restriction estimate so that it applies to functions which are *supported on a sparse union of balls of constant radius.* The idea is to exploit the estimate (6.6), i.e., that the Fourier transforms of functions which are widely separated from each other in physical space, are quasiorthogonal to each other. For completeness, recall estimate (6.6), namely

$$| < \hat{f}_0|_S, \hat{f}_1|_S >_{L^2(S,d\sigma)} | \lesssim R^{-\rho} \|f_0\|_{L^1(B(x_0,R))} \|f_1\|_{L^1(B(x_1,R))}\,. \tag{A.1}$$

Let us make these considerations now more precise by defining what we mean by *sparse collections of balls.*

Definition A.2. A collection $\{B(x_i,R)\}_{i=1}^N$ of R-balls is *sparse* if the centers x_i are $R^C N^C$ separated.

The observation (A.1) then leads to the following restriction estimate for functions supported on a sparse collection of R-balls.

Lemma A.3. *Suppose* $R_S(p \to p; \alpha)$ *holds for some* $\alpha > 0$ *and* $1 < p < 2$*. Then*

$$\|\hat{f}|_S\|_{L^p(S,d\sigma)} \lesssim R^\alpha \|f\|_{L^p(\mathbb{R}^d)}$$

whenever $\operatorname{supp} f \subseteq \bigcup_i B(x_i,R)$ *where* $\{B(x_i,R)\}$ *is a sparse collection of* R*-balls.*

The proof of this lemma will be given at the end of this subsection. Let us now continue with the proof of Theorem A.1. Suppose that $R_S(p \to p; \alpha)$ holds for some $p < 2$ and $\alpha > 0$. By the trivial (L^1, L^∞) restriction estimate, Hölder's inequality, and Marcinkiewicz interpolation (see also Remark A.5), it suffices to prove the Lorentz space estimate

$$\|\hat{f}|_S\|_{L^p(S,d\sigma)} \lesssim \|f\|_{q_0,1}\,, \tag{A.2}$$

where

$$\frac{1}{q_0} = \frac{1}{p} + \frac{A_\rho}{\log(1/\alpha)}\,.$$

and $L^{p,q}$ are the Lorentz spaces (which are Banach spaces, see, e.g., Stein and Weiss [169, Chapter V, §3, Theorem 3.22]) which are equipped with the norm

$$\|g\|_{L^{p,q}(X,d\mu)} := p^{1/q} \left\| t\mu\{|g|>t\}^{1/p} \right\|_{L^q(\mathbb{R}_+,dt/t)} \quad \text{where } 1 < p \leq \infty \text{ and } 1 \leq q \leq \infty\,.$$

(Note that the $L^{p,q}$ spaces can also be defined for $0 < p \leq 1$ and $0 < q < 1$; however, they are not Banach spaces anymore, as they cannot be normed, see also [169, Chapter 5, §5.12]).

By averaging over translations, it suffices to show (A.2) when f is a measure supported on a discrete lattice $\mathbb{Z}^d$ and the $L^{q_0,1}$ norm is replaced by the discrete norm $\ell^{q_0,1}$. One may then replace f by $f * \chi$ (and come back to the continuous norm on $L^{q_0,1}$) where χ is the characteristic function of the cube of size c, and $c \sim 1$ is chosen such that $\hat{\chi}$ is positive on the unit sphere. Combining these two reductions we see that it suffices to verify (A.2) when f is constant on c-cubes.

Since we are working in $L^{q_0,1}$, we may take $f = \mathbf{1}_E$ for some set E which we can assume to be the union of c-cubes. Thus, we are left to prove

$$\|\widehat{\mathbf{1}_E}|_S\|_{L^p(S,d\sigma)} \le A_\alpha \|\mathbf{1}_E\|_{L^{q_0,1}} \sim A_\alpha |E|^{\frac{1}{p}+\frac{A_\rho}{\log(1/\alpha)}} . \tag{A.3}$$

This will be accomplished with the help of the following Calderón–Zygmund type lemma which covers such a set E by a reasonably small number of sparse collections of balls where one has some modest control on the size of the balls[27].

Lemma A.4 ([181, Lemma 3.3], [180, Lemma 4.3])**.** *Let E be a union of c-cubes and $N \ge 1$. Then there exist $\mathcal{O}(N|E|^{1/N})$ sparse collections of balls which cover E such that the radius of the balls in each collection is of order $\mathcal{O}(|E|^{A^N})$.*

Deferring the proof of this lemma to the end of this subsection, we may now conclude the proof of the ε-removal lemma. If E is a union of c-cubes, then by Lemma A.4, one can cover E with $\mathcal{O}(N|E|^{1/N})$ sets E_j which are each the union of a sparse collection of balls of radius $\mathcal{O}(|E|^{A^N})$. By Vitali's covering lemma, one may assume $|E_j| \lesssim |E|$. Applying now Lemma A.3 to each such E_j, one obtains

$$\|\widehat{\mathbf{1}_{E_j}}|_S\|_{L^p(S,d\sigma)} \lesssim (|E|^{A^N})^\alpha |E|^{1/p} ,$$

and therefore, by the triangle inequality,

$$\|\widehat{\mathbf{1}_E}|_S\|_{L^p(S,d\sigma)} \lesssim N|E|^{1/N} (|E|^{A^N})^\alpha |E|^{1/p} .$$

Thus, (A.3) follows by taking $N = A^{-1}\log(1/\alpha)$ for a sufficiently large A.

Proof of Lemma A.3. Our first step is to modify the restriction hypothesis slightly, namely denote by $\tilde{\mathcal{R}}$ the restriction operator to the annulus A_R, i.e., a R^{-1}-neighborhood $\mathcal{N}_{1/R}(\mathbb{S}^{d-1})$ of $\mathbb{S}^{d-1}$ of thickness R^{-1} around the sphere $\mathbb{S}^{d-1}$. (Recall that we denoted the classical restriction operator by $\mathcal{R}$). The restriction hypothesis $R_S(p \to p; \alpha)$ then implies (see also Demeter [63, Proposition 1.27])

$$\|\tilde{\mathcal{R}} f\|_{L^p(\mathbb{R}^d)} = \|\hat{f}|_{A_R}\|_{L^p(\mathbb{R}^d)} \lesssim R^{-1/p+\alpha} \|f\|_{L^p(\mathbb{R}^d)} \quad \text{whenever supp } f \subseteq B(x_0,R) \tag{A.4}$$

(for any x_0 by translational symmetry) by averaging the restriction hypothesis over all $(1+\mathcal{O}(R^{-1}))$-dilations.

Now, take $f = \sum_i f_i \varphi_i$ with supp $f_i \subseteq B(x_i,R)$ (where $\{B(x_i,R)\}_{i=1}^N$ is the sparse collection of R-balls) and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ satisfies

$$\text{supp } \hat{\varphi} \subseteq B(0,1) , \quad \varphi|_{B(0,1)} \ge 0 , \quad \varphi_i(x) = \varphi\left(\frac{x-x_i}{R}\right) .$$

Since the f_i are disjointly supported, i.e., $\|f\|_p^p = \sum_i \|f_i\|_p^p$, and $\mathcal{R} f = \sum_i \tilde{\mathcal{R}} f_i * \hat{\varphi}_i\big|_{\mathbb{S}^{d-1}}$ (because $|\eta| \le |\eta - \xi| + |\xi| \le R^{-1} + |\xi|$ for $\xi \in S$ and $\xi - \eta \in \text{supp}\hat{\varphi}$ implies $\eta \in \mathcal{N}_{1/R}(\mathbb{S}^{d-1})$), the assertion follows from

$$\left\| \sum_i F_i * \hat{\varphi}_i\big|_{\mathbb{S}^{d-1}} \right\|_{L^p(\mathbb{S}^{d-1},d\sigma)} \lesssim R^{1/p} \left(\sum_i \|F_i\|_p^p \right)^{1/p} \qquad \text{for all } F_i \in L^p(\mathbb{R}^d) ,$$

---

[27]The version below is copied from [181, Lemma 3.3], while a more detailed version is contained in [180, Lemma 4.3].

taking $F_i = \tilde{\mathcal{R}} f_i$, and using the modified restriction hypothesis (A.4). This estimate follows immediately for $p = 1$, since

$$\| \sum_i F_i * \hat{\varphi}_i|_{\mathbb{S}^{d-1}} \|_1 \leq \sum_i \int_{\mathbb{S}^{d-1}} d\xi \int_{\mathbb{R}^d} d\eta \, |F_i(\eta)||\hat{\varphi}_i(\xi-\eta)| \lesssim R \sum_i \|F_i\|_1 \,.$$

By real interpolation, it therefore suffices to prove the estimate for $p = 2$. Renaming $\hat{f}_i = F_i$ and applying Plancherel's theorem, the estimate is equivalent to

$$\| \sum_i \hat{f}_i * \hat{\varphi}_i|_{\mathbb{S}^{d-1}} \|_2 = \| \sum_i \mathcal{R}(f_i \varphi_i) \|_2 \lesssim R^{1/2} \left(\sum_i \|f_i\|_2^2 \right)^{1/2} = R^{1/2} \| \{ \|f_i\|_{L^2(\mathbb{R}^d)} \}_i \|_{\ell^2} \,. \quad \text{(A.5)}$$

where we may interpret $\vec{f} = (f_1, ..., f_N)$ and $\vec{\varphi} = (\varphi_1, ..., \varphi_N)$ as elements of ℓ^2. Introduce $T : \ell^2(L^2(\mathbb{R}^d)) \to L^2(\mathbb{R}^d)$ defined by $Tf = \mathcal{R}\langle \vec{\varphi}, \vec{f} \rangle_{\ell^2}$, i.e., the left side of the last estimate equals $\|T\vec{f}\|_2$. Then $T^* : L^2(\mathbb{R}^d) \to \ell^2(L^2(\mathbb{R}^d))$ acts as $(T^*g)_j = \varphi_j \mathcal{R}^* g$ for $j \in \{1, ..., N\}$ and $g \in L^2(S)$. By Schwarz in L^2 and then in ℓ^2,

$$\begin{aligned} \| \sum_i \mathcal{R}(f_i \varphi_i) \|_{L^2}^2 &= (f, T^*Tf)_{\ell^2 L^2} = \sum_{j \in \mathbb{N}} (f_j, \varphi_j \mathcal{R}^* \mathcal{R} \sum_{i \in \mathbb{N}} \varphi_i f_i)_{L^2} \\ &\leq \sum_{j \in \mathbb{N}} \left(\|f_j\|_{L^2} \cdot \| \varphi_j \mathcal{R}^* \mathcal{R} \sum_{i \in \mathbb{N}} \varphi_i f_i \|_{L^2} \right) \\ &\leq \|f\|_{\ell^2 L^2} \left(\sum_{j \in \mathbb{N}} \| \varphi_j \mathcal{R}^* \mathcal{R} \sum_{i \in \mathbb{N}} \varphi_i f_i \|_{L^2}^2 \right)^{1/2} \end{aligned}$$

for all $f = (f_1, f_2, ...) \in \ell^2 L^2(\mathbb{R}^d)$. Thus, it suffices to prove

$$\|T^*Tf\|_{\ell^2 L^2} = \left(\sum_j \| \varphi_j \mathcal{R}^* \mathcal{R} \sum_i (\varphi_i f_i) \|_2^2 \right)^{1/2} \lesssim R \left(\sum_i \|f_i\|_2^2 \right)^{1/2} = R \|f\|_{\ell^2 L^2} \,. \quad \text{(A.6)}$$

This will follow from self-adjointness of T^*T in $\ell^2 L^2$, and the Schur test in ℓ^2 (recall Lemma 4.18)

$$\sup_j \sum_i \| \varphi_j \mathcal{R}^* \mathcal{R} \varphi_i \|_{2 \to 2} \lesssim R$$

which in turn will follow from the estimates

$$\| \varphi_i \mathcal{R}^* \mathcal{R} \varphi_i \|_{2 \to 2} \lesssim R \quad \text{(A.7a)}$$

$$\| \varphi_j \mathcal{R}^* \mathcal{R} \varphi_i \|_{2 \to 2} \lesssim (RN)^{-C} \quad \text{for } j \neq i \,. \quad \text{(A.7b)}$$

To prove the former estimate, it suffices to prove

$$\| \hat{\varphi}_i * (d\sigma(\hat{\varphi}_i * g)) \|_2 \lesssim R \|g\|_2$$

by Plancherel's theorem. This estimate, however, follows from the corresponding (obvious) $L^\infty \to L^\infty$ estimate, duality (since the operator $g \mapsto \hat{\varphi}_i * (d\sigma(\hat{\varphi}_i * g))$ is self-adjoint in $L^2(S)$) and interpolation. (That is we use that if an operator $T : L^2 \to L^2$ is self-adjoint and obeys $\|T\|_{L^\infty \to L^\infty} \leq A$ for some $A > 0$, then $\|T\|_{L^1 \to L^1} = \|T^*\|_{L^\infty \to L^\infty} = \|T\|_{L^\infty \to L^\infty} \leq A$ and so $\|T\|_{L^2 \to L^2} \leq A$ by interpolation.) Similarly, it suffices to prove the $L^\infty \to L^\infty$ analog of (A.7b) to prove (A.7b) itself. This estimate follows from the rapid decay of φ_j and φ_i for $|x_i - x_j| \gg R$ (which is the case due to the sparsity of the collection) and the decay $|\widehat{d\sigma}(x_i - x_j)| \lesssim (1 + |x_i - x_j|)^{-(d-1)/2} \lesssim (RN)^{-C}$ (for some other C) again because of the sparsity of the collection. □

Proof of Lemma A.4. For $0 \leq k \leq N$, we define radii by $R_0 = 1$ and $R_{k+1} = |E|^C R_k^C$, i.e., $R_k = \mathcal{O}(|E|^{C^K})$ for each k. (In particular, $R_1 = |E|^C$).

For $k \geq 1$, we recursively set

$$E_k := \{x \in E : \ x \notin E_j \text{ for } j < k \text{ and } |E \cap B(x, R_k)| \leq |E|^{k/N}\}$$

and note that $\bigcup_{k=1}^N E_k = E$ [28]. By construction and the hypothesis, we have for every $1 \leq k \leq N$ and $x \in E_k$,

$$|E \cap B(x, R_k)| \gtrsim |E|^{(k-1)/N} \,.$$

Thus, for every $x \in E_k$, the set $E_k \cap B(x, R_k)$ can be covered by $\mathcal{O}(|E|^{1/N})$ R_{k-1}-balls which implies that the entire set E_k can be covered by $\mathcal{O}(|E|^{1/N})$ R_{k-1}-balls which are R_k-separated. Since the cardinality of these collections can be at most $\mathcal{O}(|E|)$, the definition of R_k shows that the collections are indeed sparse what had to be shown. □

Remark A.5. Let us shortly convince ourselves that it indeed suffices to prove the Lorentz type estimate (A.2), i.e., $\|\hat{f}|_S\|_{L^p(S,d\sigma)} \lesssim \|f\|_{q_0,1}$ and the trivial estimate $\|\hat{f}|_S\|_\infty \lesssim \|f\|_1$ to deduce $\|\hat{f}|_S\|_{L^q(S,d\sigma)} \lesssim \|f\|_{L^q(\mathbb{R}^d)}$. Recall the numerology of the problem, i.e., $q < q_0 < p$ where $1/q_0 = 1/p + A_\rho/\log(1/\alpha)$. We make use of the following result, which can, e.g., be found in `[Theorem 4.6 in https://www.guillermorey.me/documents/Lorentz.pdf]` which is in fact based on Tao's notes `[Course 245C, https://terrytao.wordpress.com/2009/03/30/245c-notes-1-interpolation-of-lp-spaces/]` on interpolation of L^p spaces.

Theorem A.6 (Marcinkiewicz)**.** *Let T be a sublinear operator and suppose $0 < p_i, q_i \leq \infty$ ($i = 1, 2$) and $q_1 \neq q_2$. If T satisfies*

$$\|Tf\|_{L^{q_i,\infty}} \lesssim_i \|f\|_{L^{p_i,1}} \quad i = 1, 2$$

for all f in an appropriate dense function space, then for all $1 \leq r \leq \infty$ and $0 < \theta < 1$ such that $q_\theta > 1$, we have

$$\|Tf\|_{L^{q_\theta,r}} \lesssim_{p_1,p_2,q_1,q_2,r,\theta} \|f\|_{L^{p_\theta,r}} \,.$$

In our case, $q_1 = \infty$, $p_1 = 1$, $q_2 = p$, $p_2 = q_0 \in (q, p)$, and $p_\theta = q_\theta = r$ (since $\|f\|_{L^{p,p}} = \|f\|_p$, see also [169, p. 192]). As usual, $1/p_\theta := (1 - \theta)/p_1 + \theta/p_2$ and q_θ is defined analogously.

The condition of the former theorem is obviously satisfied for $i = 1$ (because of the trivial restriction estimate) whereas the condition for $i = 2$ follows from $L^{p,r_1} \subseteq L^{p,r_2}$ for any $0 < r_1 \leq r_2$ (see [169, Theorem 3.11]), i.e., $\|\hat{f}|_S\|_{L^{p,\infty}(S,d\sigma)} \leq \|\hat{f}|_S\|_{L^p(S,d\sigma)}$ here, and the assumed Lorentz type estimate $\|\hat{f}|_S\|_{L^p(S,d\sigma)} \lesssim \|f\|_{q_0,1}$. Finally, θ is determined by

$$\theta = \left(2 - \frac{1}{p} - \frac{A_\rho}{\log(1/\alpha)}\right)^{-1}$$

which is contained in $(0, 1)$ if α satisfies $\alpha < \exp(-A_\rho p/(p - 1))$. □

A.2. **Oscillatory integrals related to the Fourier transform.** We follow [167, Section IX.1]. Let us discuss the oscillatory integral (the *extension operator*)

$$(T_\lambda f)(x) = \int_{\mathbb{R}^{d-1}} \mathrm{e}^{i\lambda\varphi(\xi,x)} \psi(\xi, x) f(\xi) \, d\xi \,, \quad \lambda > 0 \,, \tag{A.8}$$

[28]One might imagine that, for a connected, star-shaped set E, E_1 is the union of very small sets sitting at the boundary, E_2 is the union of a bit bigger sets sitting at the inner boundary of E_1 and so on when finally only a "bubble" E_N sitting at the center is going to be left.

mapping functions on $\mathbb{R}^{d-1}$ to functions on $\mathbb{R}^d$. We simultaneously consider the dual operator (the *restriction operator*)

$$(T_\lambda^* f)(\xi) = \int_{\mathbb{R}^d} \mathrm{e}^{-i\lambda\varphi(\xi,x)}\overline{\psi}(\xi,x)f(x)\,dx\,. \tag{A.9}$$

`[Note that x and ξ are interchanged in [167] which is somewhat abusing the standard convention.]` Here, $\psi \in C_c^\infty(\mathbb{R}^{d-1}\times\mathbb{R}^d)$ is a fixed smooth function of compact support in x and y. The phase function φ is real-valued and smooth. We assume that, on the support of ψ, the phase function satisfies a non-degeneracy and a curvature condition (the Carleson–Sjölin conditions).

Let us start with the *non-degeneracy condition.* We require that for each $(\xi^0,x^0)\in \mathrm{supp}\psi \subseteq \mathbb{R}^{d-1}\times\mathbb{R}^d$, the bilinear form $B(u,v)$ on $\mathbb{R}^{d-1}\times\mathbb{R}^d$, defined by

$$B(u,v) = \langle v,\nabla_\xi\rangle\langle u,\nabla_x\rangle\varphi(\xi,x)\big|_{(\xi^0,x^0)} = \left(\sum_{j=1}^{d-1}\sum_{k=1}^{d} v_j\cdot u_k\,\frac{\partial^2\varphi(\xi,x)}{\partial\xi_j\partial x_k}\right)(\xi^0,x^0) \tag{A.10}$$

has maximal rank $d-1$ (cf. (4.7)).

As a result, there exists a (unique up to sign) vector $\overline{u}\in\mathbb{R}^d$, $|\overline{u}|=1$, so that the scalar function

$$\xi \mapsto \langle \overline{u}, \nabla_x\varphi(\xi,x^0)\rangle$$

has a critical point at $\xi=\xi^0$. Our further assumption is that this critical point is nondegenerate, i.e., we suppose that the associated $(d-1)\times(d-1)$ quadratic form is nonsingular, i.e.,

$$\det\left(\frac{\partial^2}{\partial\xi_i\partial\xi_j}\langle\overline{u},\nabla_x\varphi(\xi,x^0)\rangle\right)\neq 0 \tag{A.11}$$

at $\xi=\xi^0$. Note that this is precisely the *curvature condition* (4.8) that we imposed earlier in Theorem 4.6. The above two conditions are therefore just the *Carleson–Sjölin* conditions.

Theorem A.7. *Under the above assumptions on φ, the operator* (A.8) *satisfies the estimate*

$$\|T_\lambda f\|_{L^q(\mathbb{R}^d)} \lesssim \lambda^{-n/q}\|f\|_{L^p(\mathbb{R}^{d-1})} \tag{A.12}$$

where

$$q = \left(\frac{d+1}{d-1}\right)p' \quad \textit{and} \quad 1\le p\le 2\,.$$

Remark A.8. In several applications however, the above oscillatory integrals arise in combinations with kernels of singular integral operators. Phong and Stein [146] (see also [147]) considered the following situation. Let T be a L^2 bounded operator that is representable by a distribution kernel K, i.e., $(Tf)(x)=\int K(x,y)f(y)\,dy$ for $f\in\mathcal{S}$, where K satisfies $|\partial_y^\beta\partial_x^\alpha K(x,y)|\lesssim |x-y|^{-d-|\alpha|-|\beta|}$. Let $\varphi(x,y)$ be a real smooth phase function, let $\psi\in C_c^\infty(\mathbb{R}^d\times\mathbb{R}^d)$, and assume $\det(\partial_{x_i}\partial_{x_j}\varphi)$ has no zeros on the support of ψ. Consider the operator

$$(T_\lambda f)(x) = \int_{\mathbb{R}^d}\mathrm{e}^{i\lambda\varphi(x,y)}K(x,y)\psi(x,y)f(y)\,dy\,,$$

defined by

$$\langle g, T_\lambda f\rangle = \int_{\mathbb{R}^d}dx\int_{\mathbb{R}^d}dy\;\overline{g(x)}\mathrm{e}^{i\lambda\varphi(x,y)}K(x,y)\psi(x,y)f(y)\,.$$

Then the L^2 operator norm of T_λ remains bounded as $\lambda\to\infty$.

Proof. See Stein [167, Chapter IX, Section §1.2, Theorem 1] or [164, Theorem 10].

It suffices to prove the case $p = 2$ since the case $p = 1$ is trivial and the rest follows by interpolation. By duality, the asserted bound for $p = 2$ is equivalent to

$$\|T_\lambda^* F\|_{L^2(\mathbb{R}^{d-1})} \lesssim \lambda^{-d/r'} \|F\|_{L^r(\mathbb{R}^d)} \quad \text{for} \quad r = \frac{2(d+1)}{d+3}$$

where

$$(T_\lambda^* F)(\xi) = \int_{\mathbb{R}^d} \mathrm{e}^{-i\lambda\varphi(\xi,x)} \overline{\psi}(\xi,x) F(x)\, dx\,, \quad \xi \in \mathbb{R}^{d-1}\,.$$

Let us now rewrite the squared L^2 norm as

$$\|T_\lambda^* F\|^2_{L^2(\mathbb{R}^{d-1})} = \langle F, TT^* F\rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_\lambda(x,y) \overline{F(x)} F(y)\, dy\, dx$$

with the kernel

$$K_\lambda(x,y) = \int_{\mathbb{R}^{d-1}} \mathrm{e}^{i\lambda[\varphi(\xi,x)-\varphi(\xi,y)]} \psi(\xi,x) \overline{\psi(\xi,y)}\, d\xi\,. \tag{A.13}$$

Thus, it suffices to see that K_λ is the kernel of an $L^r(\mathbb{R}^d) \to L^{r'}(\mathbb{R}^d)$ bounded operator whose norm does not exceed a constant times $\lambda^{-2d/r'}$.

Our plan is to use Theorem 4.3 for the L^2-boundedness of *non-degenerate* oscillatory integrals (in $\mathbb{R}^d$). To apply this theorem, we construct an appropriate new phase function $\tilde{\varphi}$ on $\mathbb{R}^d \times \mathbb{R}^d$. Because of our assumptions on φ, we can construct it in such a way that it indeed satisfies the following two (non-degeneracy) conditions. Writing $\Xi = (\xi, \xi_d)$ with $\xi = (\xi_1, ..., \xi_{d-1}) \in \mathbb{R}^{d-1}$, the constructed $\tilde{\varphi}$ shall then obey

(1) $\tilde{\varphi}(\Xi, x) = \varphi(\xi, x) + \varphi_0(x)\xi_d$ and
(2) $\det(\nabla_\Xi \nabla_x \tilde{\varphi}) \neq 0$.

In fact, $\nabla_\xi \nabla_x \varphi$ already has rank $d-1$ by the non-degeneracy condition, i.e., we need only chose $\varphi_0(x)$ such that $\langle u, \nabla_x\rangle\varphi_0(x) \neq 0$ to increase the rank of the matrix $\nabla_\Xi \nabla_x \tilde{\varphi}$ to d.

Now, as in the shortest proof of the Tomas–Stein theorem (see Subsection 4.3), we construct an analytic family of kernels K_λ^s on $\mathbb{R}^d \times \mathbb{R}^d$ by setting

$$K_\lambda^s(x,y) = \frac{\mathrm{e}^{s^2}}{\Gamma(s/2)} \int_{\mathbb{R}^d} \mathrm{e}^{i\lambda[\tilde{\varphi}(\Xi,x)-\tilde{\varphi}(\Xi,y)]} \psi(\xi,x) \overline{\psi(\xi,y)} |\xi_d|^{-1+s} \nu(\xi_d)\, d\Xi$$

with $d\Xi = d\xi\, d\xi_d$ and where $\nu \in C_c^\infty(\mathbb{R})$ is a bump function at the origin. Let $\mathcal{T}_\lambda^s$ be the associated integral operator. By an integration by parts, setting $s = 0$, and applying the fundamental theorem of calculus along with $\tilde{\varphi}(\Xi, x)|_{\Xi=(\xi,0)} = \varphi(\xi, x)$, we have

$$K_\lambda^0(x,y) = K_\lambda(x,y)\,. \tag{A.14}$$

Remember that we want to estimate $\|\mathcal{T}_\lambda^0\|_{r\to r'} \lesssim \lambda^{-2d/r'}$ via complex interpolation. Next, by Theorem 4.3 for non-degenerate oscillatory integral operators, we have $\|T_\lambda^{1+it}\|_{2\to 2} \lesssim \lambda^{-d/2}$ for all $t \in \mathbb{R}$ because of the non-degeneracy condition on $\tilde{\varphi}$. Finally, we claim the following $L^1 \to L^\infty$ estimate, namely

$$|K_\lambda^{-(d-1)/2+it}(x,y)| \lesssim 1\,. \tag{A.15}$$

To see this, recall $\tilde{\varphi}(\Xi, x) = \varphi(\xi, x) + \varphi_0(x)\xi_d$ and write

$$K_\lambda^s(x,y) = K_\lambda(x,y) \cdot \tilde{\nu}_s(\lambda(\varphi_0(y) - \varphi_0(x)))$$

where

$$\tilde{\nu}_s(\lambda(\varphi_0(y) - \varphi_0(x))) = \frac{\mathrm{e}^{s^2}}{\Gamma(s/2)} \int_{-\infty}^{\infty} \mathrm{e}^{i\xi_d \cdot \lambda(\varphi_0(y)-\varphi_0(x))} \nu(\xi_d) |\xi_d|^{-1+s}\, d\xi_d\,.$$

Since

$$|\hat{\nu}_{-(d-1)/2}(\lambda(\varphi_0(y)-\varphi_0(x)))| \lesssim |\lambda(\varphi_0(y)-\varphi_0(x))|^{\frac{d-1}{2}} \leq \|\nabla\varphi_0\|_\infty^{\frac{d-1}{2}} \cdot (\lambda|x-y|)^{\frac{d-1}{2}}$$

for large arguments (i.e., large λ), we are left to show

$$|K_\lambda(x,y)| \lesssim (\lambda|x-y|)^{-(d-1)/2}\,.$$

In proving this, we may assume that the integrand is supported in a sufficiently small neighborhood around some $\xi=\xi^0$ (for otherwise we can write it as a finite sum of such integrals). Then, we observe that $\varphi(\xi,x)-\varphi(\xi,y)=\nabla_x\varphi(\xi,z)\cdot(x-y)+\mathcal{O}(|x-y|^2)$. So, the claimed bound on $|K_\lambda(x,y)|$ just follows from the estimates for non-degenerate oscillatory integrals in $d-1$ dimensions (Theorem 4.3) because of the non-degeneracy condition for φ which clearly still holds when we freeze one variable (see also the remark before Theorem 4.6). (In fact, if $x-y$ does not point in the "critical direction" of u, which arises in the non-degeneracy condition, we even get $|K_\lambda(x,y)|\lesssim(\lambda|x-y|)^{-N}$ for any $N\in\mathbb{N}$ since we can integrate by parts as often as we wish.) This concludes the proof of the theorem. □

Bourgain [16] proved that the theorem can in fact not be improved beyond the range $1\leq p\leq 2$ when $d\geq 3$. To see this, let $d=3$. Then there is an appropriate φ and a bounded f having compact support such that

$$\|T_\lambda f\|_q \gtrsim \lambda^{-1/2-1/q} \quad \text{as} \quad \lambda\to\infty\,.$$

This is however only consistent with the assertion of Theorem A.7 if $q\geq 4$ (i.e., $p\leq 2$). To prove this lower bound, take

$$\varphi(\xi,x)=\xi\cdot x'+\frac{1}{2}\langle A(x_3)\xi,\xi\rangle$$

for $\xi\in\mathbb{R}^2$, $x=(x',x_3)\in\mathbb{R}^3$ and $A(x_3)$ is a real, symmetric 2×2 matrix, depending smoothly on x_3. We will now impose two conditions on $x_3\mapsto A(x_3)$.

(1) $dA(x_3)/dx_3$ is invertible for each x_3. This condition guarantees the *curvature condition* at the critical point, namely that the $(d-1)\times(d-1)$ quadratic form satisfies $\det(\partial_{x_i}\partial_{x_j}\langle\overline{u},\nabla_x\varphi(\xi,x^0)\rangle)\neq 0$ at $\xi=\xi^0$.

(2) $\operatorname{rank}(A(x_3))\equiv 1$ for all x_3, i.e., the non-degeneracy condition is satisfied.

It is easy to check that these two conditions are compatible, and that indeed there are smooth functions $x_3\mapsto A(x_3)$ that satisfy both simultaneously. Now let $f(x)=1$ on the support of ψ. Then

$$(T_\lambda f)(x)=\int_{\mathbb{R}^2} \mathrm{e}^{i\lambda\varphi(\xi,x)}\psi(\xi)\,d\xi\,.$$

Let $S=\{x\in\mathbb{R}^3: x'\in\operatorname{Ran}(A(x_3))\}$. In view of our assumptions on $\operatorname{rank}(A(x_3))$, we see that S is a smooth hypersurface. Note that if $x\in S$, the quadratic function $\xi\mapsto\varphi(\xi,x)$ has a critical point, and moreover the rank of $\partial_{x_i}\partial_{x_j}\varphi(\xi,x)$ is exactly 1. Thus if $x\in S$, we can show, using stationary phase, that

$$|(T_\lambda f)(x)|\sim\lambda^{-1/2} \quad \text{as} \quad \lambda\to\infty\,.$$

The estimate also holds in a tubular neighborhood of S whose radius is a small multiple of λ^{-1}. The result is that

$$\|T_\lambda f\|_q \gtrsim \lambda^{-1/2}\lambda^{-1/q}\,,$$

and the result is proved.

See also Bourgain [16] where it is also shown that for a certain class of phases φ, one does have

$$\|T_\lambda f\|_q \lesssim \lambda^{-d/q}\|f\|_\infty$$

for some q with $q<2(d+1)/(d-1)$.

Appendix B. Reverse Littlewood–Paley inequality for slabs in $d = 2$

In this section we will describe an argument of Córdoba and Fefferman [50] (see also Córdoba [51, 52]) yielding the reverse Littlewood–Paley inequality (8.1) in $d = 2$ if $p = 4$. Combining this with the analysis of Section 8 yields a full proof of the restriction conjecture in this case.

The heart of the argument is the fact that the Minkowski sums of all pairs of slabs $\theta + \theta' = \{\xi + \xi' : \xi \in \theta, \xi' \in \theta'\}$ have only bounded overlap (which in turn is somewhat a consequence of the fact that two circles in $\mathbb{R}^2$ intersect in at most two points).

Proposition B.1. *Let f be a smooth function with* $\operatorname{supp} \hat{f} \subseteq \mathcal{N}_{1/R}(\mathbb{P}^1)$*. With the notation of Section 8, the inequality*

$$\|f\|_{L^4(\mathbb{R}^2)} \lesssim \left\| \left(\sum_{\theta : R^{-1/2}-\text{slab}} |f_\theta|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^2)}$$

holds.

Proof. By the Fourier support condition, we have

$$\|f\|_4^4 = \||f|^2\|_2^2 \sim \int_{\mathbb{R}^2} \left| \sum_\theta f_\theta(x) \right|^2 \left| \sum_\theta \overline{f_{\theta'}(x)} \right|^2 dx = \left\| \sum_{\theta,\theta'} f_\theta \overline{f_{\theta'}} \right\|_2^2 \tag{B.1}$$

We distinguish now between the cases $\operatorname{dist}(\theta,\theta') \lesseqgtr \text{const}\, R^{-1/2}$ and start with $\operatorname{dist}(\theta,\theta') \lesssim R^{-1/2}$. By Cauchy–Schwarz,

$$\left| \sum_{\theta,\theta' : \operatorname{dist}(\theta,\theta') \lesssim R^{-1/2}} f_\theta \overline{f_{\theta'}} \right| \le 2 \sum_\theta |f_\theta|^2 \sum_{\theta' : \operatorname{dist}(\theta,\theta') \lesssim R^{-1/2}} 1 \lesssim \sum_\theta |f_\theta|^2 \,,$$

i.e., it suffices to estimate the right side of (B.1) where the summation is restricted to slabs which are at least $R^{-1/2}$-separated. In particular, it suffices to show

$$\left\| \sum_{\theta,\theta' : \operatorname{dist}(\theta,\theta') \gtrsim R^{-1/2}} f_\theta \overline{f_{\theta'}} \right\|_2^2 \lesssim \sum_{\theta,\theta' : R^{-1/2}-\text{slab}} \|f_\theta f_{\theta'}\|_2^2$$

which can be interpreted as the statement that $f_\theta \overline{f_{\theta'}}$ are pairwise almost orthogonal. (Observe that this right side just agrees with the right side of the statement of the proposition).

Observing that the left side of the claimed inequality equals

$$\left\| \sum_{\theta,\theta' : \operatorname{dist}(\theta,\theta') \gtrsim R^{-1/2}} \hat{f}_\theta * \overline{\hat{f}_{\theta'}} \right\|_2^2$$

by Plancherel's theorem and that

$$\operatorname{supp} \hat{f}_\theta * \overline{\hat{f}_{\theta'}} \subseteq \theta - \theta' \,,$$

it suffices to prove that the number of overlaps of $\theta - \theta'$ is bounded, i.e.,

$$\left| \left\{ \theta, \theta' : R^{-1/2} - \text{slab} : \operatorname{dist}(\theta,\theta') \gtrsim R^{-1/2} \text{ and } \xi \in \theta - \theta' \right\} \right| \lesssim 1 \quad \text{for all } \xi \in \mathbb{R}^2 \,.$$

To prove this, consider the pairs θ_1, θ_1' and θ_2, θ_2' which are such that $\theta_1 - \theta_1' \cap \theta_2 - \theta_2' \neq \emptyset$ and $\operatorname{dist}(\theta_j, \theta_j') \gtrsim R^{-1/2}$ (for $j = 1, 2$). In particular, that means that there are $y_j \in \theta_j$ and

$y_j' \in \theta_j'$ such that $y_1 - y_1' = y_2 - y_2'$. Moreover, since θ_j and θ_j' belong to $\mathcal{N}_{1/R}(\mathbb{P}^1)$, there are $t_j, t_j' \in [0,1]^{d-1}$ such that

$$|y_j - (t_j, t_j^2)| \lesssim R^{-1} \quad \text{and} \quad |y_j' - (t_j', (t_j')^2)| \lesssim R^{-1} \quad \text{for } j = 1, 2\,.$$

Defining $\underline{t}_j = (t_j, t_j^2)$, adding and subtracting $(y_1 - y_1') - (y_2 - y_2') = 0$, and using the above estimate yields

$$|(\underline{t}_1 - \underline{t}_1') - (\underline{t}_2 - \underline{t}_2')| \lesssim R^{-1}$$

which means in particular

$$|(t_1 - t_1') - (t_2 - t_2')| \lesssim R^{-1} \quad \text{and} \quad |(t_1^2 - (t_1')^2) - (t_2^2 - (t_2')^2)| \lesssim R^{-1}\,.$$

From these estimates, it can be inferred `[by expanding everything?]`

$$|t_1 - t_1'| \cdot |(t_1 + t_1') - (t_2 + t_2')| \lesssim R^{-1}\,.$$

Since $\mathrm{dist}(\theta_1, \theta_1') \gtrsim R^{-1/2}$, it follows that

$$|(t_1 + t_1') - (t_2 + t_2')| \lesssim R^{-1/2}$$

and in particular

$$|t_1 - t_2| \lesssim R^{-1/2} \quad \text{and} \quad |t_1' - t_2'| \lesssim R^{-1/2}$$
$$\Rightarrow |y_1 - y_2| \lesssim R^{-1/2} \quad \text{and} \quad |y_1' - y_2'| \lesssim R^{-1/2}\,.$$

But that means that for a given pair θ_1, θ_1' there are only $\mathcal{O}(1)$ choices of pairs θ_2, θ_2' such that $\theta_1 - \theta_1' \cap \theta_2 - \theta_2' \neq \emptyset$ which means

$$\left|\left\{\theta, \theta' : R^{-1/2} - \text{slab} : \mathrm{dist}(\theta, \theta') \gtrsim R^{-1/2} \text{ and } \xi \in \theta - \theta'\right\}\right| \lesssim 1 \quad \text{for all } \xi \in \mathbb{R}^2$$

as asserted. □

Appendix C. Interpolation theorems

See, e.g., Tao's notes on harmonic analysis or Grafakos [96, Section 1.4].

C.1. **Repetition on Lorentz spaces.** See, e.g., Folland [79, Section 6.4], Adams–Fournier [1, pp. 221], the notes by G. Rey `https://www.guillermorey.me/documents/Lorentz.pdf`, Grafakos [96, Section 1.4], Triebel [191], and Bennett and Sharpley [5]. For interpolation theory, consider Bennett–Sharpley (once more) and in particular Bergh and Löfström [8].

Let (X, σ, μ) be a measure space, i.e., a set X equipped with a σ-algebra of subsets of it and a function μ from the σ-algebra to $[0, \infty]$ that satisfies $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} \mu(B_j)$$

for any sequence B_j of pairwise disjoint elements of the σ-algebra. The function μ is called a (positive) measure on X and elements of the σ-algebra of X are called *measurable sets.*

Definition C.1. Let f be a measurable function on X. Its *distribution function* $\lambda_f : \mathbb{R}_+ \to [0, \infty]$ is defined by

$$\lambda_f(\alpha) := \mu\left(\{x \in X : |f(x)| > \alpha\}\right)$$

We collect some classic properties, see, e.g., Grafakos [96, Propositions 1.1.3 and 1.1.4].

Proposition C.2.

(1) λ_f is non-increasing and right-continuous.
(2) If $|f| \leq |g|$, then $\lambda_f \leq \lambda_g$.

(3) *If* $|f_n|$ *increases to* $|f|$*, then* λ_{f_n} *increases to* λ_f*.*
(4) *If* $f = g + h$*, then* $\lambda_f(\alpha + \beta) \le \lambda_g(\alpha) + \lambda_h(\beta)$*.*
(5) *We have the layer cake representation*

$$f(x) = \int_0^\infty \mathbf{1}_{\lambda_f(\alpha)}(x)\,d\alpha = \int_0^\infty \mathbf{1}_{[0,f(x)]}(\alpha)\,d\alpha\,.$$

(6) *We have*

$$\int_X |f(x)|^p\,dx = p\int_0^\infty \alpha^p \lambda_f(\alpha)\,\frac{d\alpha}{\alpha}\,.$$

(7) *We have* $\|f\|_\infty = \inf\{\alpha \ge 0 : \lambda_f(\alpha) = 0\}$*.*

Chebyshev's inequality asserts

$$\lambda_f(\alpha) \le \alpha^{-p}\|f\|_p^p$$

which leads to the definition of weak L^p spaces.

Definition C.3. Let $0 < p < \infty$. Then we denote by $L^{p,\infty}(X)$ the class of all functions whose quasi-norm (i.e., the triangle inequality only holds up to some constant)

$$\begin{aligned}\|f\|_{p,\infty}^p &:= \sup_{\alpha>0} \alpha^p \lambda_f(\alpha)\\ &= \inf\left\{C > 0 :\ \lambda_f(\alpha) \le \alpha^{-p}C \quad \text{for all } \alpha > 0\right\}\end{aligned}$$

is finite.

Remarks C.4.
(1) Check that both definitions actually coincide!
(2) By Chebyshev's inequality, we immediately see $\|f\|_{p,\infty} \le \|f\|_p$, i.e., $L^p \subseteq L^{p,\infty}$.
(3) By construction, $L^{\infty,\infty}$ isometrically coincides with L^∞.
(4) Since $\lambda_{g+h}(\alpha) \le \lambda_g(\alpha/2) + \lambda_h(\alpha/2)$, it is easy to see that $\|g + h\|_{p,\infty} \le 2C_p(\|g\|_{p,\infty} + \|h\|_{p,\infty})$.

Example C.5. Let $p \in [1,\infty)$ and $f(x) = |x|^{-d/p}$, then $f \notin L^p(\mathbb{R}^d)$ for any p, but $f \in L^{p,\infty}$ since

$$|\{x \in \mathbb{R}^d : |x|^{-d/p} > \alpha\}| = \int\limits_{|x|<\alpha^{-p/d}} dx \ \sim \alpha^{-p}\,.$$

The equimeasurable decreasing rearrangement of f is the function f^* on $[0,\infty)$, defined by

$$f^*(t) := \inf_{\alpha>0}\{\lambda_f(\alpha) \le t\} = \inf_{\alpha\ge0}\{\lambda_f(\alpha) \le t\}\,,$$

which is a non-increasing function since λ_f is non-increasing. In particular, $\lambda_{f^*}(\alpha) = \lambda_f(\alpha)$. Let us now define the Lorentz quasi-norm.

Definition C.6. Let f be a measurable function on X and $0 < p, q \le \infty$. We define the Lorentz quasi-norm as

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

By definition, $L^{p,p}$ coincides isometrically with L^p.

Proposition C.7. *Let* f *be a measurable function on* X *and* $0 < p, q \le \infty$*. Then*

$$\|f\|_{p,q} = \begin{cases} p^{1/q}\left(\int_0^\infty \left(\alpha\lambda_f(\alpha)^{1/p}\right)^q \ \frac{d\alpha}{\alpha}\right)^{1/q} & \textit{if } q < \infty\,,\\ \sup_{\alpha>0} \alpha\lambda_f(\alpha)^{1/p} & \textit{if } q = \infty\,.\end{cases}$$

Proof. See Grafakos [96, Proposition 1.4.9]. □

We collect some useful properties.

Lemma C.8 (Monotone convergence). *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions with $|f_n| \nearrow |f|$ almost everywhere. Then $\|f\|_{p,q} = \lim_{n\to\infty} \|f_n\|_{p,q}$.*

Proof. See Lemma 1.3 in Rey's notes. □

Lemma C.9 (Fatou). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions. Then*

$$\| \liminf_{n\to\infty} f_n\|_{p,q} \leq \liminf_{n\to\infty} \|f\|_{p,q} \, .$$

Proof. See Lemma 1.9 in Rey's notes. □

Theorem C.10. *Let $0 < p < \infty$ and $0 < q \leq \infty$, then $L^{p,q}$ is a quasi Banach space, i.e., it is complete and satisfies the quasi triangle inequality. For $p, q > 1$, they are normable and in particular actual Banach spaces.*

Proof. See Grafakos [96, Theorem 1.4.11]. □

Proposition C.11 (Nestedness). *Let $0 < p \leq \infty$ and $0 < q < r \leq \infty$. Then $\|f\|_{p,r} \lesssim_{p,q,r} \|f\|_q$, i.e., $L^{p,q} \subseteq L^{p,r}$.*

Proof. See Grafakos [96, Proposition 1.4.10]. □

Proposition C.12 (Hölder's inequality). *Let $0 < p_1, p_2, p < \infty$ and $0 < q_1, q_2, q \leq \infty$ obey*

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} \quad \textit{and} \quad \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \, .$$

Then $\|fg\|_{p,q} \lesssim \|f\|_{p_1,q_1} \|g\|_{p_2,q_2}$.

Proof. See Tao [186, Lecture 1, Theorem 6.9]. □

More details concerning the following proposition can be found in Grafakos [96].

Proposition C.13 (Dual characterization). *Let $1 < p < \infty$ and $1 \leq q \leq \infty$. Then for any $f \in L^{p,q}$,*

$$\|f\|_{p,q} \sim_{p,q} \sup \left\{ \left| \int \overline{g(x)} f(x) \, d\mu(x) \right| \|g\|_{p',q'} \leq 1 \right\} \, .$$

Proof. See Tao [186, Lecture 1, Theorem 6.12]. □

C.2. **Marcinkiewicz interpolation.** Typically, the Marcinkiewicz interpolation theorem is stated under the condition that an operator satisfies two weak-type estimates. Recall that if X and Y are two measure spaces and T is a linear operator from functions of X to functions of Y, then T is said to be of strong-type (p, q) if

$$\|Tf\|_{L^q(Y)} \lesssim \|f\|_{L^p(X)} \quad \text{for all } f \in L^p(X) \, .$$

We say that T is of weak-type (p, q) if

$$|\{y \in Y : |(Tf)(y)| \geq \lambda\}| \lesssim \|f\|_p^q \lambda^{-q} \quad \text{for all } \lambda > 0 \, , f \in L^p(X) \, .$$

Clearly, the strong-type estimate implies the weak-type estimate. One can weaken this concept even further by only considering functions f which are characteristic functions of a set. This leads to the notion of *restricted weak-type estimates.* We say that T is of restricted weak type (p, q) if

$$|\{y \in Y : |(T\mathbf{1}_E)(y)| \geq \lambda\}| \lesssim |E|^{q/p} \lambda^{-q} \quad \text{for all } \lambda > 0 \, , E \subseteq X \, . \tag{C.1}$$

Note that every characteristic function belongs to $L^{p,1}$ with

$$\|\mathbf{1}_E\|_{p,1} = \text{const } |E|^{1/p}$$

The enhanced Marcinkiewicz interpolation theorem (see, e.g., Tao's notes [179, Lecture 2, Lemma 2.3] or Grafakos [96, Theorem 1.4.19] and Tao [186, Lecture 1, Lemma 8.5]) therefore says that if T is $L^{p_j,1} \to L^{q_j,\infty}$ bounded for $j \in \{0,1\}$, then T is $L^{p_\theta,r} \to L^{q_\theta,r}$ bounded for all $0 < r \leq \infty$.

Remark C.14. There is also a result by Bourgain [25] (see also Grafakos [96, p. 71]), but see also Carbery et al [40, Section 6.2] saying that if the sequence of linear operators T_j maps

$$\|T_j\|_{A_i \to B_i} \lesssim_i 2^{j\alpha_i}$$

for $i \in \{0,1\}$ and normed vector spaces A_i, B_i with $\alpha_0 < 0 < \alpha_1$, then $T = \sum_j T_j$ extends to a bounded operator mapping $\overline{A}_{\theta,1}$ to $\overline{B}_{\theta,\infty}$. (Recall $\overline{A} = (A_0, A_1)$ and $\overline{A}_{\theta,1}$ and $\overline{B}_{\theta,\infty}$ are the Lions–Peetre interpolation spaces.) A more precise and explicit version is formulated in the following proposition.

Proposition C.15 (Bourgain interpolation (Grafakos' version [96]))**.** *Let $0 < p_0 < p_1 < \infty$ and $0 < \beta_0, \beta_1, M_0, M_1 < \infty$. Suppose that for $k \in \mathbb{Z}$ a family of sublinear operators $\{T_k\}$ is of restricted weak-type (p_0, p_0) with constant $M_0 2^{-k\beta_0}$ and of restricted weak-type (p_1, p_1) with constant $M_1 2^{k\beta_1}$ for all $k \in \mathbb{Z}$. Then there is a constant $C = C(\beta_0, \beta_1, p_0, p_1)$ such that $\sum_{k\in\mathbb{Z}} T_k$ is of restricted weak-type (p,p) with constant $CM_0^{1-\theta}M_1^\theta$ where $\theta = \beta_0/(\beta_0+\beta_1)$ and $p^{-1} = (1-\theta)/p_0 + \theta/p_1$.*

Proposition C.16 (Bourgain interpolation (Carbery et al version [40]))**.** *Let $0 < p_0, p_1, q_0, q_1 < \infty$ and $0 < \beta_0, \beta_1, M_0, M_1 < \infty$. Suppose that for $k \in \mathbb{Z}$ a family of sublinear operators $\{T_k\}$ satisfies*

$$\|T_j\|_{L^{p_0}\to L^{q_0}} \leq M_0 2^{-\beta_0 j} \quad \textit{and} \quad \|T_j\|_{L^{p_1}\to L^{q_1}} \leq M_1 2^{+\beta_1 j}\,.$$

Then there is a constant $C = C(\beta_0, \beta_1, p_0, p_1, q_0, q_1)$ such that

$$\|\sum_{k\in\mathbb{Z}} T_k f\|_{L^{q,\infty}} \leq C\, M_0^{1-\theta} M_1^\theta \|f\|_{L^{p,1}}$$

where $\theta = \beta_0/(\beta_0+\beta_1)$ and p and q are as usual.

It is convenient to reformulate (C.1) in a more symmetric, dual formulation.

Lemma C.17. *Let $1 < p, q < \infty$. Then, one has* (C.1) *if and only if*

$$|\langle \mathbf{1}_F, T\mathbf{1}_E\rangle| \lesssim |E|^{1/p}|F|^{1/q'} \tag{C.2}$$

for all $E \subseteq X$ and $F \subseteq Y$.

This should be compared to the dual strong-type estimate

$$|\langle g, Tf\rangle| \lesssim \|f\|_p \|g\|_{q'}\,.$$

Proof. For our purposes, we only need the implication (C.1)⇒(C.2). (To prove the reverse direction, one sets $F = \{\mathrm{Re}(T\mathbf{1}_E) > \lambda\}$.) Using the triangle inequality, the layer cake representation, and Fubini to do the x-integration first, we have

$$\begin{aligned} |\langle \mathbf{1}_F, T\mathbf{1}_E\rangle| &\leq \int_F |(T\mathbf{1}_E)(x)|\,dx = \int_F \int_0^\infty \mathbf{1}_{\{|T\mathbf{1}_E|>\lambda\}}(\lambda)\,d\lambda\,dx \\ &= \int_0^\infty |\{x \in F : |(T\mathbf{1}_E)(x)| > \lambda\}|\,d\lambda\,. \end{aligned}$$

We have two estimates for the integrand. The first is just $|F|$. The second is $\mathcal{O}(|E|^{q/p}\lambda^{-q})$ by assumption. Thus, the integral can be estimated by

$$\mathcal{O}\left(\int_0^\infty \min\{|F|, |E|^{q/p}\lambda^{-q}\}\, d\lambda\right)$$

which yields the assertion after an elementary calculation. □

C.3. **Stein interpolation.** See Stein and Weiss [169, Chapter V, Theorem 4.1].

Appendix D. Some remarks on the uncertainty principle

D.1. **Bernstein inequalities.** We follow the nice exposition of Wolff [200, Chapter 5] and the survey of Folland and Sitaram [80]. For the following discussion it will be helpful to remember that for an invertible linear map $T : \mathbb{R}^d \to \mathbb{R}^d$, one has

$$\widehat{f \circ T} = |\det(T)|^{-1} \hat{f} \circ T^{-t}$$

where T^{-t} denotes the inverse transpose of T.

For us, most of the time, the uncertainty principle is the following heuristic statement. *If a measure μ is supported on an ellipsoid E, then for many purposes $\hat{\mu}$ may be regarded as being constant on any dual ellipsoid E^*.*

The simplest rigorous statement is as follows.

Proposition D.1 (L^2 Bernstein inequality)**.** *Assume that $f \in L^2$ and* $\operatorname{supp} \hat{f} \subseteq B_0(R)$ *for some $R > 0$. Then f is C^∞ and it holds that*

$$\|D^\alpha f\|_2 \leq (2\pi R)^{|\alpha|} \|f\|_2\,.$$

Proof. Since $\hat{f}$ is compactly supported, f is in fact holomorphic and the claimed estimate just follows from Plancherel. □

A corresponding statement is also true in L^p, but proving this and other related results needs a different argument (namely, the Mikhlin–Hörmander theorem) since there is no Plancherel theorem. In this context we state somme lemmas that are helpful to construct compactly supported functions in Fourier space from Schwartz functions in physical space.

Lemma D.2. *There is a fixed Schwartz function φ such that if $f \in L^1 + L^2$ and* $\operatorname{supp} \hat{f} \subseteq B_0(R)$*, then*

$$f = \varphi^{R^{-1}} * f$$

where $\varphi^{R^{-1}}(x) = R^d \varphi(Rx)$.

Proof. Take $\varphi \in \mathcal{S}$ such that $\hat{\varphi}|_{B_0(1)} = 1$, i.e., $\widehat{\varphi^{R^{-1}}}|_{B_0(R)} = 1$. Thus, $(\varphi^{R^{-1}} * f - f)^\wedge \equiv 0$ which shows the assertion. □

Lemma D.3. *There are radial bump functions $\hat{\chi}$ that satisfy $\chi \geq 0$ and $\chi > \mathbf{1}_{B_0(1)}$.*

Proof. If g is an even bump function, then take $\hat{\chi}(\xi) = A^d B(g * g)(A\xi)$ for some $A, B > 0$. □

Lemma D.4. *There exists a radial $0 < \varphi \in \mathcal{S}(\mathbb{R}^d)$ such that* $\operatorname{supp} \hat{\varphi} \subseteq (-1/2, 1/2)^d$ *and with the property that*

$$\sum_{n \in \mathbb{Z}^d} \varphi(x - n) = 1\,, \quad x \in \mathbb{R}\,.$$

Proof. See Schlag–Shubin–Wolff [152, Lemmas 2.4 and Lemma 3.1]. We only present the proof for $d = 1$. The proof for higher dimensions is almost identical.

In Fourier space the claimed partition of unity reads

$$\hat{\varphi}(\xi) \sum_{n\in\mathbb{Z}} \mathrm{e}^{-2\pi i n\xi} = \sum_{k\in\mathbb{Z}} \hat{\varphi}(k)\delta_k(\xi) = \delta_0(\xi) \tag{D.1}$$

where the first equality follows from Poisson summation ($\sum_n f(n) = \sum_k \hat{f}(k)$). To ensure the second equality, it suffices to take $\operatorname{supp}\hat{\varphi} \subseteq (-1/2, 1/2)$ and set $\hat{\varphi}(0) = 1$[29]. To obtain the positivity, start with any even Schwartz function φ_0 with $\operatorname{supp}\widehat{\varphi_0} \subseteq (-1/4, 1/4)$ and $\widehat{\varphi_0}(0) = 1$. Since φ_0^2 extends to an entire function on $\mathbb{C}$, one has

$$\mathrm{mes}[\varphi_0^2 = 0] = 0\,.$$

Therefore, $\varphi = \varphi_0^2 * \varphi_0^2 > 0$ everywhere, whereas

$$\hat{\varphi} = [\widehat{\varphi_0} * \widehat{\varphi_0}]^2$$

has support in $(-1/2, 1/2)$. Finally observe that

$$\hat{\varphi}(0) = \left(\int \widehat{\varphi_0}(\xi)\widehat{\varphi_0}(-\xi)\, d\xi\right)^2 = \left(\int \widehat{\varphi_0}(\xi)^2\, d\xi\right)^2 > 0\,.$$

The second equation in (D.1) uses that $\widehat{\varphi_0}$ is even whereas positivity follows since $\widehat{\varphi_0}$ is real. Hence,

$$\sum_{n\in\mathbb{Z}} \varphi(x-n) = \hat{\varphi}(0)\,, \quad x \in \mathbb{R}$$

by the preceeding argument. Dividing by the right-hand side finishes the proof. □

Proposition D.5 (L^p Bernstein inequality)**.** *Suppose that* $f \in L^1 + L^2$ *and* $\operatorname{supp}\hat{f} \subseteq B_0(R)$. *Then the following assertions hold.*

(1) For any α *and* $p \in [1,\infty]$,

$$\|D^\alpha f\|_p \lesssim R^{|\alpha|}\|f\|_p\,.$$

(2) For any $1 \le p \le q \le \infty$,

$$\|f\|_q \lesssim R^{d(\frac{1}{p}-\frac{1}{q})}\|f\|_p\,.$$

With the help of the second assertion it becomes obvious that Bernstein inequalities are an invaluable tool in the analysis of (nonlinear) PDEs. The inequalities say that, for localized frequency, low Lebesgue integrability can be upgraded to higher integrability (i.e., smoothness) at the cost of certain powers of N. In fact, this cost is a gain when the frequency is small.

Proof. As before, let $\psi = \varphi^{R^{-1}}$ such that $f = \psi * f$. Then the first claim just follows from

$$\|\nabla\psi\|_1 = \|\varphi\|_1 \cdot R$$

and Young's inequality. To prove the second assertion, we note

$$\|\psi\|_r = \|\varphi\|_r \cdot R^{d/r'}$$

for any $r \in [1,\infty]$. Thus, for r being defined by $1 + 1/q = 1/p + 1/r$, Young's inequality yields

$$\|f\|_q = \|\psi * f\|_q \le \|\psi\|_r \|f\|_p \lesssim R^{d/r'}\|f\|_p = R^{d(\frac{1}{p}-\frac{1}{q})}\|f\|_p\,,$$

thereby showing the second claim. □

[29]One could have also obtained this directly since $\sum_{n\in\mathbb{Z}} \mathrm{e}^{-2\pi i n\xi} = \delta_{\xi,0}$.

With this warm-up, we are ready to extend the above $L^p \to L^q$ bounds to ellipsoids instead of balls using change of variables. An *ellipsoid* is a set of the form

$$E = \{x \in \mathbb{R}^d : \sum_j \frac{|(x-a)\cdot e_j|^2}{r_j^2} \leq 1\} \tag{D.2}$$

for some $a \in \mathbb{R}^d$ (the center of E), some choice of orthonormal basis vectors $\{e_j\}$ (the axes), and some choice of positive numbers r_j (the axis lengths). We define the *dual ellipsoid* E^* to E as the ellipsoid having the same axes as E but with reciprocal axis lengths, i.e., if E is given by (D.2), then E^* should be of the form

$$\{x \in \mathbb{R}^d : \sum_j r_j^2 |(x-b)\cdot e_j|^2 \leq 1\} \tag{D.3}$$

for some choice of the center point b.

Proposition D.6 (L^p Bernstein inequality for an ellipsoid)**.** *Suppose that* $f \in L^1 + L^2$ *and* $\operatorname{supp} \hat{f} \subseteq E$ *for some ellipsoid* E*. Then*

$$\|f\|_q \lesssim |E|^{\frac{1}{p}-\frac{1}{q}} \|f\|_p$$

if $1 \leq p \leq q \leq \infty$.

This statement reflects the heuristic fact that faster decay of the Fourier transform (i.e., the smaller the ellipsoid E is) yields better smoothness properties (in terms of integrability) of the function.

One could similarly extend the gradient bounds of the previous statements to ellipsoids centered at the origin, but that statement is awkward since one has to weight different directions differently, so we ignore this here.

Proof. Let k be the center of E and T be a linear map taking the unit ball onto $E - k$. Let $S = T^{-t}$ be the inverse transpose of T, i.e., also $T = S^{-t}$. Let furthermore $f_1(x) = \mathrm{e}^{-2\pi i k\cdot x} f(x)$ and $g = f_1 \circ S$. Since $\widehat{f \circ T} = |\det(T)|^{-1} \hat{f} \circ T^{-t}$, we have

$$\hat{g}(\xi) = |\det S|^{-1} \hat{f}_1(S^{-t}(\xi)) = |\det S|^{-1} \hat{f}(S^{-t}(\xi + k)) = |\det T| \hat{f}(T(\xi + k)) \,.$$

Thus, $\hat{g}$ is supported in the unit ball, so by the L^p Bernstein inequality for balls, $\|g\|_q \lesssim \|g\|_p$. On the other hand,

$$\|g\|_q = |\det S|^{-1/q} \|f\|_q = |\det T|^{1/q} \|f\|_q = |E|^{1/q} \|f\|_q$$

and likewise with q replaced by p. So

$$|E|^{1/q} \|f\|_q \lesssim |E|^{1/p} \|f\|_p$$

as claimed. □

D.2. **Locally constant lemma.** Finally, we will also prove a "pointwise statement", roughly saying that if $\operatorname{supp} \hat{f} \subseteq E$ for some ellipsoid E, then f is *roughly constant* on the dual ellipsoid E^* (and rapidly decreasing away from it if $\hat{f}$ is assumed to be smooth). In fact, we shall show that the values of f on E^* can *morally(!)* be controlled by the average over E^*.

To formulate this precisely, let N be a large number and let $\varphi(x) = (1 + |x|^2)^{-N}$. Suppose an ellipsoid E^* is given. Then define $\varphi_{E^*}(x) = \varphi(T(x - k))$, where k is the center of E^* and T is a self-adjoint linear map taking $E^* - k$ onto the unit ball. If T_1 and T_2 are two such maps, then

$T_1 \circ T_2^{-1}$ is an orthogonal transformation, so φ_{E^*} is well defined. Essentially φ_{E^*} roughly equals 1 on E^* and decays rapidly as one moves away from E^*. We could also write more explicitly

$$\varphi_{E^*}(x) = \left(1 + \sum_j \frac{|(x-k)\cdot e_j|^2}{r_j^2}\right)^{-N} .$$

Proposition D.7 (Locally constant lemma)**.** *Suppose that* $f \in L^1 + L^2$ *and* supp $\hat{f} \subseteq E$ *for some ellipsoid* E*. Then for any dual ellipsoid* E^* *and any* $z \in E^*$*,*

$$|f(z)| \lesssim_N |E^*|^{-1} \int_{\mathbb{R}^d} |f(x)| \varphi_{E^*}(x)\, dx \equiv |E^*|^{-1} \|f\|_{L^1(\varphi_{E^*}(x)dx)} . \tag{D.4}$$

Proof. Assume first that E is the unit ball so that E^* is also the unit ball. Then f is the convolution of itself with a fixed Schwartz function ψ. Accordingly,

$$\begin{aligned} |f(z)| &\leq \int_{\mathbb{R}^d} |f(x)||\psi(x-z)|\, dx \lesssim_N \int_{\mathbb{R}^d} |f(x)|(1+|x-z|^2)^{-N}\, dx \\ &\lesssim_N \int_{\mathbb{R}^d} |f(x)|(1+|x|^2)^{-N}\, dx \end{aligned}$$

where we used the rapid decay of ψ and $1 + |x-z|^2 \gtrsim 1 + |x|^2$ uniformly in $|x|$ when $|z| \leq 1$. This proves the assertion when $E = E^*$ is the unit ball.

Next, suppose E is centered at zero but E and E^* are otherwise arbitrary. Let k and T be as above (T took $E^* - k$ to the unit ball, i.e., T^t maps the unit ball to E, and T^{-1} maps E onto the unit ball, i.e., T^{-t} maps the unit ball onto $E^* - k$; more precisely, these maps also take any translate of one set to the according translate of the other!), and consider

$$g(x) = f(T^{-1}x + k) \quad \Leftrightarrow \quad \hat{g}(\xi) = |\det T|\ \mathrm{e}^{2\pi i k\cdot\xi} \hat{f}(T^t\xi) .$$

Thus, $\hat{g}$ is supported on $T^{-1}E$, i.e., a unit ball. According to our above findings for the unit ball, we have

$$|g(y)| \leq \int_{\mathbb{R}^d} \varphi(x)|g(x)|\, dx$$

if y belongs to *any* unit ball. Hence, it follows that

$$\begin{aligned} f(T^{-1}z + k) &\leq \int_{\mathbb{R}^d} \varphi(x)|f(T^{-1}x+k)|\, dx = |\det T| \int_{\mathbb{R}^d} \varphi_{E^*+k}(x)|f(x+k)|\, dx \\ &= |E^*|^{-1} \int_{\mathbb{R}^d} \varphi_{E^*}(x)|f(x)|\, dx \end{aligned}$$

by a change of variables and the fact $|\det T| = |E^*|^{-1}$. This shows the assertion since the above estimate holds for z in some unit ball which we may identify with $T(E^* - k)$ (since T took $E^* - k$ to the unit ball) which however means that the argument $T^{-1}z + k$ belongs to E^*. □

When E is not centered at zero, one merely needs to replace $f(x)$ by $\mathrm{e}^{-2\pi i k\cdot x} f(x)$ where k is the center of E.

Remark D.8.

(1) The above proposition is an example of an estimate "with Schwartz tails". It is not possible to make the stronger conclusion that, say, $|f(x)|$ is bounded by the average of f *over the double of* E^* when $x \in E^*$ (even in the one dimensional case with $E = E^*$ being the unit interval); i.e., taking the *average over* $\mathbb{R}^d$ is necessary! To see this, consider

a fixed Schwartz function g with $g(0) \neq 0$ whose Fourier transform is supported in the "unit interval" $[-1,1]$. Consider also the functions

$$f_N(x) = \left(1 - \frac{x^2}{4}\right)^N g(x)\,.$$

Since $\hat{f}_N$ are linear combinations of $\hat{g}$ and its derivatives, they have the same support as $\hat{g}$. Moreover, they converge pointwise boundedly to zero on $[-2,2]$, except at the origin. It follows that there can be no estimate of the value of f_N at the origin by its average over $[-2,2]$.

(2) All the estimates related to Bernstein's inequality are sharp except for the values of the constants. For instance, if E is an ellipsoid, E^* a dual ellipsoid, and $N < \infty$, then there is a function f with supp $\hat{f} \subseteq E^*$ and with

$$\|f\|_1 \geq |E|\,,$$
$$|f(x)| \leq A\varphi_E(x)\,,$$

where $\varphi_E = \varphi_E^{(N)}$ was defined above. In the case $E = E^*$ being the unit ball, this is obvious; take f to be any Schwartz function with Fourier support in the unit ball and with the appropriate L^1 norm. The general case then follows as above by making a change of variables.

(3) The name "locally constant lemma" is motivated by the following counterexample. Consider $f \in \mathcal{S}(\mathbb{R} : \mathbb{R})$ with supp$\hat{f} \subseteq [0,1]$. Then one might wonder whether f could not actually look like a sequence of peaks whose distance to each other is extremely small. The locally constant lemma says that this cannot occur. On the one hand, due to the pointwise bound $\|f\|_\infty \lesssim \|f\|_{L^1(\varphi_{E^*}\,dx)}$, one sees that the peaks must not be too big. However, these peaks then cannot add up to the given L^1 norm. Hence, f cannot be a sequence of narrow peaks, but must actually be roughly constant on the dual interval (which is just $[0,1]$ again).

The last two estimates imply that $\|f\|_p \sim |E|^{1/p}$ for any p which shows that the last proposition is also sharp.

D.3. **Localization and discretization.** The following is taken from [58]. The essence of the following results is the uncertainty principle in the following verbal form.

Localization in x-space on the scale R induces a smoothing in ξ-space on the scale R^{-1}. This amounts to discretize ξ-space on the scale R^{-1}. The roles of x and ξ are interchangable.

D.3.1. *Localization in momentum space.* Denote by $\mathcal{Q}_h$ the collection of all cubes Q_h of sidelength h. Define the weight function

$$w_{Q_h}(x) = (1 + h^{-1}\mathrm{dist}(x, Q_h))^{-100d}, \quad x \in \mathbb{R}^d, \quad Q_h \in \mathcal{Q}_h. \tag{D.5}$$

We start by restating the locally constant lemma (Proposition D.7)

Lemma D.9. *Let $v \in \mathcal{S}(\mathbb{R}^d)$ and assume that $\hat{v}$ is supported in $B(0,1/h)$. Then v is locally constant on all cubes Q_h of sidelength h in the sense that*

$$\|v\|_{L^\infty(Q_h)} \lesssim |Q_h|^{-1}\|v\|_{L^1(w_{Q_h})}.$$

This lemma allows us to compare the $L^p(\mathbb{R}^d)$ norm of a function with Fourier support contained in a ball of radius $1/h$ with its ℓ^p-norm, when sampled on a lattice of h-distant points.

Lemma D.10. *Let $v \in \mathcal{S}(\mathbb{R}^d)$ and assume that $\hat{v}$ is supported in $B(0, 1/h)$. Let $\Lambda_h \subset \mathbb{R}^d$ be a set of h-separated points. Then for any $p \geq 1$, we have*

$$\|v\|_{\ell^p(\Lambda_h)} \lesssim h^{-d/p}\|v\|_{L^p(\mathbb{R}^d)}.$$

Proof. Again by scaling, we can assume $h = 1$. Thus, let $\Lambda \subset \mathbb{R}^d$ be a set of 1-separated points. Pick a collection of cubes Q of sidelength one that cover Λ. By Lemma D.9,

$$\|v\|^p_{\ell^p(\Lambda)} = \sum_{\nu \in \Lambda} |v(\nu)|^p \lesssim \sum_Q \|v\|^p_{L^1(w_Q)},$$

Write $v = \sum_{Q'} v_{Q'}$, where $v_{Q'}$ is supported on Q'. Then

$$\|v_{Q'}\|_{L^1(w_Q)} \leq (1 + \operatorname{dist}(Q, Q'))^{-100d}\|v_{Q'}\|_{L^1(\mathbb{R}^d)}.$$

By Hölder, $\|v_{Q'}\|_{L^1(\mathbb{R}^d)} \leq \|v_{Q'}\|_{L^p(\mathbb{R}^d)}$. Hence, (by Young)

$$\sum_Q \|v\|^p_{L^1(w_Q)} \lesssim \sum_{Q,Q'} (1 + \operatorname{dist}(Q, Q'))^{-100dp}\|v_{Q'}\|^p_{L^p(\mathbb{R}^d)} \lesssim \|v\|^p_{L^p(\mathbb{R}^d)},$$

where we summed a geometric series in Q. □

D.3.2. *Discrete Fourier extension operator.* Let $M_\lambda = \{\xi \in \mathbb{R}^d : |\xi| = \lambda\}$, and consider the extension operator

$$\mathcal{E}_\lambda : L^2(M_\lambda, \mathrm{d}\sigma_\lambda) \to L^\infty(\mathbb{R}^d), \quad (\mathcal{E}_\lambda g)(x) = (g\mathrm{d}\sigma_\lambda)^\vee(x),$$

where σ_λ is surface measure on M_λ. We write $\mathcal{E} \equiv \mathcal{E}_1$ and $M \equiv M_1$, $\sigma \equiv \sigma_1$.

In the following situation we will exploit the effects of a simultaneous position and frequency (almost) localization. In local Fourier restriction theory we saw that a position localizationon scale R allowed a smoothing, and thereby a discretization, of ξ-space on scale R^{-1}.

Definition D.11. Let $\mathrm{Discres}(M, p, 2)$ be the best constant such that the following hold for each $R \geq 2$, each collection Λ^*_R consisting of $1/R$-separated points on M, each sequence $a_\nu \subset \mathbb{C}$, each ball B_R and each collection Λ_1 of 1-separated points in $\mathbb{R}^d$:

$$\|\sum_{\nu \in \Lambda^*_R} a_\nu \mathrm{e}(\nu \cdot x)\|_{\ell^{p'}(\Lambda_1 \cap B_R)} \leq \mathrm{Discres}(M, p, 2) R^{\frac{d-1}{2}} \|a_\nu\|_{\ell^2(\Lambda^*_R)}. \tag{D.6}$$

The following proposition replaces the $L^p(B(R))$-norm encountered in local Fourier restriction theory by a $\ell^p(B(R))$-norm, where we sample over points whose distance is dictated by the Fourier length scale. Thus, if we were dealing with a sphere of radius $1/h$, the sampling would have to occur over h-distant points in x-space.

Proposition D.12. *If $1 \leq p \leq \infty$, then*

$$\mathrm{Discres}(M, p, 2) \lesssim \|\mathcal{E}\|_{L^2(M,\mathrm{d}\sigma) \to L^{p'}(\mathbb{R}^d)}. \tag{D.7}$$

Moreover, if $p \geq 2$, then the reverse inequality also holds.

Proof. The claim is a special case of [63, Prop. 1.29], with one small difference. There, $\mathrm{Discres}(M, p, 2)$ is defined with the $L^{p'}(B_R)$ norm in the left hand side of (D.6). Thus, let $\mathrm{Discres}'(M, p, 2)$ be the best constant in the inequality

$$\|\sum_{\nu \in \Lambda^*_R} a_\nu \mathrm{e}(\nu \cdot x)\|_{L^{p'}(B_R)} \leq \mathrm{Discres}'(M, p, 2) R^{\frac{d-1}{2}} \|a_\nu\|_{\ell^2(\Lambda^*_R)}. \tag{D.8}$$

Then [63, Prop. 1.29] asserts that the proposition holds with $\mathrm{Discres}'(M, p, 2)$ in place of $\mathrm{Discres}(M, p, 2)$. Thus, (D.7) follows once we show that

$$\mathrm{Discres}'(M, p, 2) \gtrsim \mathrm{Discres}(M, p, 2). \tag{D.9}$$

Without loss of generality we may assume that $B_R = B(0,R)$. If we set

$$f(x) = \sum_{\nu\in\Lambda_R^*} a_\nu \mathrm{e}(\nu\cdot x), \quad \text{then} \quad \mathcal{F}(f\hat{\varphi}_R)(\xi) = \sum_{\nu\in\Lambda_R^*} a_\nu \varphi_R(\xi+\nu),$$

where φ_R is as before and $\mathcal{F}$ denotes the Fourier transform. Note that $\mathcal{F}(f\hat{\varphi}_R) = \hat{f} * \varphi_R$ is supported in an $1/R$-neighborhood of M. In particular, it is supported on the ball $B(0,2)$. Thus, for any collection Λ_1 of 1-separated points in $\mathbb{R}^d$,

$$\|f\|_{\ell^{p'}(\Lambda_1\cap B_R)} \leq \|f\hat{\varphi}_R\|_{\ell^{p'}(\Lambda_1)} \lesssim \|f\hat{\varphi}_R\|_{L^{p'}(\mathbb{R}^d)},$$

where we used $\hat{\varphi}_R \geq \mathbf{1}_{B_R}$ in the first inequality and Lemma D.10 in the second. By a partition of unity and a sparsification argument we may assume that f is supported on a disjoint union of balls of radius R. By the rapid decay of $\hat{\varphi}_R$ and by the definition of $\mathrm{Discres}'(M,p,2)$,

$$\|\hat{\varphi}_R f\|_{L^{p'}(\mathbb{R}^d)} \lesssim_N \sum_{j=1}^\infty j^{-N}\|f\|_{L^{p'}(B(x_j,R))} \lesssim \mathrm{Discres}'(M,p,2) R^{\frac{d-1}{2}} \|a_\nu\|_{\ell^2(\Lambda_R^*)},$$

where we used that (D.8) holds uniformly in the centers of the balls. Combining the last two estimates yields (D.9).

To prove the reverse inequality to (D.7), we may assume that $B_R = B(0,R)$. By [63, Prop. 1.29] it suffices to prove the reverse inequality to (D.9). Let Λ_1 be a 1-net of points $x_j \in B_R$. Let $f(x)$ be defined as above. Without loss of generality we may assume that f is supported on a disjoint collection of balls $B(x_j,10)$. Then

$$\begin{aligned}\|f\|_{L^{p'}(B_R)} &= (\sum_j \|f\|_{L^{p'}(B(x_j,10))}^{p'})^{1/p'} = (\int_{B(0,10)} \sum_j |f(x_j+y)|^{p'} \mathrm{d}y)^{1/p'} \\ &\lesssim \mathrm{Discres}(M,p,2) R^{\frac{d-1}{2}} \|a_\nu\|_{\ell^2(\Lambda_R^*)},\end{aligned}$$

where we used that (D.6) holds for each collection $x_j + y$ of 1-separated points, uniformly in y. □

D.3.3. *Stein–Tomas theorem.* The following is an immediate consequence of the Stein–Tomas theorem and Proposition D.12 (see also [63, Cor. 1.30]).

Proposition D.13. *Let* $p' \geq 2(d+1)/(d-1)$. *Then* $\mathrm{Discres}(M,p,2) \lesssim 1$.

D.4. **Preliminaries for the wave packet decomposition.** If $\hat{f}$ was smooth and real-valued and supported on some ellipsoid E, then (by an integration by parts argument, say) $f \in \mathcal{S}(\mathbb{R}^d)$ is concentrated on E^* with center of mass at the origin. In general, when $\hat{f}$ is complex-valued, one should expand $\hat{f}$ in a Fourier series where one samples at the centers of masses of all E^* tiling $\mathbb{R}^d$. Let us make this more precise and assume for simplicity that $\operatorname{supp}\hat{f} \subseteq \theta_\omega$ where $\theta_\omega \subseteq \mathbb{R}^d$ is a rectangle centered at the origin with side lengths $R^{-1/2}\times\cdots\times\mathbb{R}^{-1/2}\times R^{-1}$, oriented along $\omega \in \mathbb{S}^{d-1}$. To make the computations more accessible, let $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subseteq [-1/2,1/2]^d$ and

$$\begin{aligned}&T : \theta_\omega \to [-1/2,1/2]^d \\ &\quad T = D\circ R, \quad R\in SO(d), \quad D = \operatorname{diag}(R^{1/2},...,R^{1/2},R)\end{aligned}$$

where $R \in SO(d)$ rotates θ_ω to θ_{e_d}, and D scales the rectangle to the unit box $[-1/2,1/2]^d$. Let us denote by T_ω^a a dual rectangle with side lengths $R^{1/2}\times\cdots\times R^{1/2}\times R$, oriented along $\omega \in \mathbb{S}^{d-1}$, and centered at $a \in \mathbb{R}^d$. Let us collect all centers of masses of these dual rectangles that tile $\mathbb{R}^d$ by $\mathbb{T}_\omega$. Then

$$\hat{f} = \varphi \circ T$$

and we wish to expand

$$\hat{f}(\xi) = \sum_{a\in\mathbb{T}_\omega} f_\omega(a)\mathrm{e}^{2\pi i a\cdot\xi}\mathbf{1}_{\theta_\omega}(\xi)$$

for certain Fourier coefficients $f_\omega(a)$ that we shall now compute. We prepare for $a \in \mathbb{T}_\omega$,

$$\begin{aligned}\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle &= \int d\xi\,(\varphi\circ T)(\xi)\mathrm{e}^{2\pi i a\cdot\xi} = \frac{1}{|\det(T)|}\int d\xi\,\varphi(\xi)\mathrm{e}^{2\pi i\langle T^{-t}a,\xi\rangle}\\ &= \frac{1}{|\det(T)|}\langle \mathrm{e}^{2\pi i\langle T^{-t}a,\cdot\rangle},\varphi\rangle\,.\end{aligned}$$

Since $T^{-t}a \in \mathbb{Z}^d$ whenever $a \in \mathbb{T}_\omega$, we obtain by summing the equality over $a \in \mathbb{T}_\omega$,

$$\sum_{a\in\mathbb{T}_\omega} |\det(T)|\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle \mathrm{e}^{2\pi i\langle T^{-t}a,\xi\rangle} = \varphi(\xi)\,.$$

Hence, replacing $\xi \mapsto T\xi$, we obtain

$$\hat{f}(\xi) = (\varphi\circ T)(\xi) = |\det(T)|\sum_{a\in\mathbb{T}_\omega}\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle \mathrm{e}^{2\pi i a\cdot\xi}\mathbf{1}_{\theta_\omega}(\xi)\,. \tag{D.10}$$

Taking the inverse Fourier transform, we observe

$$f(x) \sim \sum_{a\in T_\omega}\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle \mathrm{e}^{2\pi i a\cdot\xi}\chi_{T_\omega}(x-a) \tag{D.11}$$

where (recall $|\det T| = R^{(d+1)/2}$) $\chi_{T_\omega} := R^{(d+1)/2}\mathcal{F}(\mathbf{1}_{\theta_\omega})$ which is focused on T_ω and obeys $\|\chi_{T_\omega}\|_\infty \sim 1$. Finally observe that we do not immediately get a Plancherel identity, but

$$\begin{aligned}\|\hat{f}\|_2^2 &= |\det(T)|^2\sum_{a,b\in\mathbb{T}_\omega}\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle\cdot\langle \hat{f}, \mathrm{e}^{2\pi i\langle b,\cdot\rangle}\rangle\int d\xi\,\mathrm{e}^{2\pi i\xi\cdot(a-b)}|\mathbf{1}_{\theta_\omega}(\xi)|^2(\delta_{a,b}+1-\delta_{a,b})\\ &= |\det(T)|\sum_{a\in\mathbb{T}_\omega}|\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle|^2 + \sum_{a\neq b\in\mathbb{T}_\omega}\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle\langle \hat{f}, \mathrm{e}^{2\pi i\langle b,\cdot\rangle}\rangle\int_{\mathbb{R}^d}\chi_{T_\omega}(a-b-y)\chi_{T_\omega}(y)\,dy\,.\end{aligned} \tag{D.12}$$

Since $(\chi_{T_\omega} * \chi_{T_\omega})(a-b)$ is a Schwartz function adapted to, say, a doubly dilated tube T_ω with $\|\chi_{T_\omega} * \chi_{T_\omega}\|_\infty \sim R^{(d+1)/2}$, we have (with the abbreviation $f_\omega(a) := \langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle$) for any $N \in \mathbb{N}$,

$$\begin{aligned}\sum_{a\neq b\in\mathbb{T}_\omega}|f_\omega(a)||f_\omega(b)|(\chi_{T_\omega} * \chi_{T_\omega})(a-b) &\sim_N R^{\frac{d+1}{2}}\sum_{a\neq b\in\mathbb{T}_\omega}|f_\omega(a)||f_\omega(b)|(1+|a-b|)^{-N}\\ &\lesssim_N R^{\frac{d+1}{2}}\sum_{a\in\mathbb{T}_\omega}|\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, f\rangle|^2\,.\end{aligned} \tag{D.13}$$

Therefore,

$$\|\hat{f}\|_2^2 \sim R^{(d+1)/2}\sum_{a\in\mathbb{T}_\omega}|\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, \hat{f}\rangle|^2 \tag{D.14}$$

which reflects the almost orthogonality of the $\langle \mathrm{e}^{2\pi i\langle a,\cdot\rangle}, f\rangle\chi_{T_\omega}$ for different $a \in \mathbb{T}_\omega$.

D.5. **Wave packet decomposition for the truncated paraboloid.** The following is taken from Demeter [63, Chapter 2].

Let $\Upsilon \in C_c^\infty([-4,4]^{d-1})$ with

$$\sum_{j\in\mathbb{Z}^d} \Upsilon(\xi - j) = 1\,.$$

Now let us refine the mesh a bit. Let $R \gg 1$ and rescale the lattice $\mathbb{Z}^{d-1}$ to $R^{-1/2}\mathbb{Z}^{d-1}$. Then

$$\sum_{|j|\lesssim R^{1/2}} \Upsilon(R^{1/2}\xi - j) \sim \mathbf{1}_{|\xi|\lesssim 1}$$

where $\Upsilon(R^{1/2}\xi - j)$ equals roughly an indicator function on a cube $\omega := c_\omega + [-R^{-1/2}, R^{-1/2}]^{d-1}$ with $c_\omega = R^{-1/2}j$. The c_ω denote the centers of those cubes which are roughly $R^{-1/2}$ distant from each other and overlap at most $O(1)$ many times. We collect these center of masses in the set $\Omega_R \subseteq R^{-1/2}\mathbb{Z}^{d-1}$. Then, as in (D.10) it is reasonable to decompose f in a Fourier series

$$f(\xi) = \sum_{c_\omega\in\Omega_R}\sum_{c_q\in Q_R} R^{\frac{d-1}{2}} \langle \mathrm{e}^{2\pi i\langle c_q,\cdot\rangle}, f\mathbf{1}_\omega\rangle \mathrm{e}^{2\pi i c_q\cdot\xi}\Upsilon(R^{1/2}(\xi - c_\omega)) \tag{D.15}$$

$$\equiv \sum_{c_\omega\in\Omega_R}\sum_{c_q\in Q_R} \langle \mathrm{e}^{2\pi i\langle c_q,\cdot\rangle}, f\mathbf{1}_\omega\rangle \mathrm{e}^{2\pi i c_q\cdot c_\omega}\Upsilon_{q,\omega}(\xi) \tag{D.16}$$

where $c_q \in R^{1/2}\mathbb{Z}^{d-1}$ are centers of dual cubes $c_q + [-R^{1/2}, R^{1/2}]^{d-1}$ which are $R^{1/2}$-separated and collected in the set Q_R, and

$$\Upsilon_{q,\omega}(\xi) = R^{\frac{d-1}{2}} \mathrm{e}^{2\pi i c_q\cdot(\xi - c_\omega)}\Upsilon(R^{1/2}(\xi - c_\omega))\,, \qquad \|\Upsilon_{q,\omega}\|_2^2 \sim R^{\frac{d-1}{2}}\,. \tag{D.17}$$

For future use, let us record the following almost orthogonality property, valid for any weight $w_{q,\omega} \in \mathbb{C}$ (such as $w_{q,\omega} = w_q\delta_{\omega,\omega_0}$ for instance),

$$\|\sum_{c_\omega\in\Omega_R}\sum_{c_q\in Q_R} w_{q,\omega}\Upsilon_{q,\omega}\|_2 \sim R^{\frac{d-1}{4}} \left(\sum_{c_\omega\in\Omega_R}\sum_{c_q\in Q_R} |w_{q,\omega}|^2\right)^{1/2} \tag{D.18}$$

which is a consequence of the fact that the cubes ω are only finitely overlapping (for the $\sum_{c_\omega\in\Omega_R}$ summation) and similar computations as in (D.12)-(D.14) (for the $\sum_{c_q\in Q_R}$ summation). Our goal is to understand $F_S^* f$ in the case of the truncated paraboloid $S = \mathbb{P}^{d-1} = \{(\xi,\xi^2) : \xi \in [-1,1]^{d-1}\}$. To that end, we first record

$$(F_S^*\Upsilon_{q,\omega})(x) = \mathrm{e}^{-2\pi i x\cdot(c_\omega, c_\omega^2)} \int \Upsilon_{q,\omega}(\eta)\mathrm{e}^{2\pi i\varphi_{x,q,\omega}(\eta)}\,d\eta$$

with a complicated expression for the phase (which just comes from exploiting the galilean symmetries of $\mathbb{P}^{d-1}$)

$$\varphi_{x,q,\omega}(\eta) = \eta\cdot\frac{x' - c_q + 2c_\omega x_d}{R^{1/2}} + \eta^2\frac{x_d}{R}\,.$$

By stationary phase, we anticipate that $F_S^*\Upsilon_{q,\omega}$ will be concentrated on a $R^{1/2}\times\cdots\times R^{1/2}\times R$ tube centered at $(c_q, 0)$ pointing in the direction $(-2c_\omega, 1)$ (which just follows from the observation that the old $x' = 0$ point gets mapped to the new point x' which satisfies $x' + 2c_\omega x_d = 0$). That is, $F_S^*\Upsilon_{q,\omega}(x)$ decays rapidly whenever (x', x_d) is no critical point in the sense that $\nabla_\eta\varphi_{x,q,\omega} = 0$. More precisely, recall Theorem 20.14 which says

$$WF(F_S^*\Upsilon_{q,\omega}) = \{(x, \nabla_x\varphi_{x,q,\omega}(\eta)) : (x,\eta) \in \mathbb{R}^d\setminus\{0\}\times\mathrm{supp}(\Upsilon_{q,\omega}),\ \nabla_\eta\varphi_{x,q,\omega}(\eta) = 0\}\,.$$

That is, the singularities will propagate along rays pointing in the direction $\nabla_x\varphi_{x,q,\omega}$. To make these statements more precise we introduce the following definitions.

Definition D.14 (Tubes and wave packets)**.** (1) We denote by $T_{q,\omega}$ the spatial tube in $\mathbb{R}^d$ given by

$$T_{q,\omega} = \{x = (x', x_d) \in \mathbb{R}^d : |x' - c_q + 2c_\omega x_d| \le R^{1/2}, |x_d| < R\} \tag{D.19}$$

The collection of these tubes for fixed ω is denoted by $\mathbb{T}_\omega$. The collection of all tubes is denoted by $\mathbb{T}$.

(2) For $M \ge 1$, let

$$MT_{q,\omega} = \{x = (x', x_d) \in \mathbb{R}^d : |x' - c_q + 2c_\omega x_d| \le MR^{1/2}, |x_d| \le R\} \tag{D.20}$$

denote the dilate of $T_{q,\omega}$ around its central axis.

(3) For each tube $T = T_{q,\omega}$ we write $\Upsilon_{q,\omega} \equiv \Upsilon_T$ and $F_S^* \Upsilon_T \equiv \varphi_T$. The latter function (or any scalar multiple thereof) is called wave packet.

The following theorem (see [63, Theorem 2.2]) summarizes the main features of the wave packet decomposition.

Theorem D.15 (Wave packet decomposition)**.** *Let $f \in C^\infty([-4,4]^{d-1})$, then there is a decomposition*

$$f = \sum_{T \in \mathbb{T}} f_T \tag{D.21}$$

with $\operatorname{supp} f_T \subseteq \omega_T$ *for some* $\omega_T = c_{\omega,T} + [-R^{-1/2}, R^{-1/2}]^{d-1}$ *with* $c_{\omega,T} \in \Omega_R$. *Let* $F_S^* f_T = a_T \varphi_T$ *with* $a_T \in \mathbb{C}$ *so that*

$$F_S^* f = \sum_{T \in \mathbb{T}} a_T \varphi_T \,. \tag{D.22}$$

Then the φ_T obey for any $k \ge 1$,

$$\|\varphi_T\|_\infty \lesssim 1\,, \qquad \|\varphi_T\|_2 \lesssim R^{(d+1)/4} \tag{D.23}$$

$$\|\varphi_T\|_{L^\infty(\mathbb{R}^{d-1} \times [-R,R] \setminus MT)} \lesssim_k M^{-k}\,, \quad M \ge 1 \tag{D.24}$$

$$\operatorname{supp} \hat{\varphi}_T \subseteq \{(\xi, \xi^2) : \xi \in \omega_T\}\,, \tag{D.25}$$

the a_T obey the Plancherel similarity (recall $\|\Upsilon_T\|_2^2 \sim R^{\frac{d-1}{2}}$)

$$\|f\|^2 \sim R^{\frac{d-1}{2}} \sum_{T \in \mathbb{T}} |a_T|^2 \tag{D.26}$$

$$\|f \mathbf{1}_\omega\|_2^2 \sim R^{\frac{d-1}{2}} \sum_{T \in \mathbb{T}_\omega} |a_T|^2\,, \tag{D.27}$$

and the coefficients f_T obey the Plancherel similarity

$$\|f\|_2^2 \sim \sum_{T \in \mathbb{T}} \|f_T\|_2^2\,. \tag{D.28}$$

In particular, th choices

$$a_T = \mathrm{e}^{2\pi i c_q \cdot c_\omega} \cdot \langle \mathrm{e}^{2\pi i \langle c_q, \cdot \rangle}, f \mathbf{1}_\omega \rangle \quad \text{and} \quad f_T = a_T \Upsilon_T$$

are admissible.

Proof. Taking a_T and f_T as above, then the bound

$$\|\varphi_T\|_{L^\infty(\mathbb{R}^{d-1} \times [-R,R] \setminus MT)} \lesssim_k M^{-k}\,, \quad M \ge 1$$

follows from non-stationary phase arguments since

$$\inf_{\eta \in \operatorname{supp}(\Upsilon),\, x \in (\mathbb{R}^{d-1} \times [-R,R]) \setminus MT} |\nabla_\eta \varphi_{x,q,\omega}(\eta)| \gtrsim M\,.$$

The bound $\|\varphi_T\|_{L^\infty(\mathbb{R}^d)} \lesssim 1$ is immediate while

$$\|f\mathbf{1}_\omega\|_2^2 \sim R^{\frac{d-1}{2}} \sum_{T\in\mathbb{T}_\omega} |a_T|^2\,,$$

follows from the almost orthogonality of the $\Upsilon_{q,\omega}$ (D.18) or Parseval's identity

$$\sum_{c_q\in Q_R} |\langle \mathrm{e}^{2\pi i\langle c_q,\cdot\rangle}, f\rangle|^2 = R^{-\frac{d-1}{2}} \|f\mathbf{1}_\omega\|_2^2\,.$$

Summing this over all $\omega \in \Omega_R$ and using the fact that these cubes overlap at most $O(1)$ many times (to exploit almost orthogonality), one infers

$$\|f\|^2 \sim R^{\frac{d-1}{2}} \sum_{T\in\mathbb{T}} |a_T|^2\,.$$

Note also that

$$\|f\|_2^2 \sim \sum_{T\in\mathbb{T}} \|f_T\|_2^2\,.$$

follows from the bound $\|\Upsilon_T\|_2^2 \sim R^{(d-1)/2}$ and

$$\operatorname{supp}\hat\varphi_T \subseteq \{(\xi,\xi^2) : \xi\in\omega_T\}\,,$$

which in turn follows from a direct computation. Note that $\hat\varphi_T$ is a measure supported on a hypersurface. □

D.6. **Other interesting uncertainty principles.** Another interesting variant of the uncertainty principle was found by Shubin–Vakilian–Wolff [153, Theorem 2.1].

Definition D.16. Let $\rho(x) = \min\{1, 1/|x|\}$. Then a set $E\subseteq\mathbb{R}^d$ is called *ε-thin* if

$$|E\cap B_x(\rho(x))| \le \varepsilon\,|B_x(\rho(x))|\,,\quad x\in\mathbb{R}^d\,.$$

Theorem D.17 ([153, Theorem 2.1])**.** *There are $\varepsilon>0$ and $C<\infty$ such that if E and F are two ε-thin sets in $\mathbb{R}^d$, then for any $f\in L^2(\mathbb{R}^d)$, it holds that*

$$\|f\|_2 \le C\left(\|f\|_{L^2(E^c)} + \|\hat f\|_{L^2(F^c)}\right)\,.$$

Clearly, the theorem says that f and $\hat f$ cannot both be concentrated on small sets at the same time. There are numerous related results in the literature, see, e.g., Fefferman [76] or Havin–Jöricke [104].

We keep track of the following lemma which says that sharp cut-offs in spatial variables automatically lead to frequency smearings on the inverse scale.

Lemma D.18. *Let $N_1, N_2 > 0$, $N > N_1+N_2$, and $F:\mathbb{R}^d\to\mathbb{R}^d$ be measurable. Let $\gamma_{1/N}(\xi) := N^d\gamma(N\xi)$ where $\check\gamma$ is a smooth bump function on $\mathbb{R}^d$ such that $\check\gamma(x)=1$ for $|x|\le 1$, i.e., $\gamma_{1/N}$ is a smoothing operator in frequency space on scale N^{-1}. Then*

$$\mathbf{1}_{|x|\le N_1} F(D)\mathbf{1}_{|x|\le N_2} = \mathbf{1}_{|x|\le N_1}\mathcal{F}^{-1}\left(F(\xi)*\gamma_{1/N}\right)\mathcal{F}\,\mathbf{1}_{|x|\le N_2}\,. \tag{D.29}$$

Analogously, for any surface measure $d\sigma$ on a codimension one manifold S that is embedded in $\mathbb{R}^d$, we have

$$\mathbf{1}_{|x|\le N_1} F_S^* F_S \mathbf{1}_{|x|\le N_2} = \mathbf{1}_{|x|\le N_1}\mathcal{F}^{-1}\left(d\sigma * \gamma_{1/N}\right)\mathcal{F}\,\mathbf{1}_{|x|\le N_2} \tag{D.30}$$

where F_S and F_S^* are the usual Fourier restriction and extension operators.

Proof. Let $f \in \mathcal{S}(\mathbb{R}^d)$, then

$$\left(\mathbf{1}_{|x|\le N_1} F(D) \mathbf{1}_{|x|\le N_2} f\right)(x) = \mathbf{1}_{|x|\le N_1} \int \check{F}(x-y) \mathbf{1}_{|y|\le N_2} f(y)\, dy\,.$$

Since $|x| \le N_1$ and $|y| \le N_2$, we automatically have $|x-y| \le N_1 + N_2 < N$. Thus, with the smooth bump function $\check{\gamma}$, we obtain

$$\begin{aligned}\left(\mathbf{1}_{|x|\le N_1} F(D) \mathbf{1}_{|x|\le N_2} f\right)(x) &= \mathbf{1}_{|x|\le N_1} \int \check{F}(x-y)\check{\gamma}(|x-y|/N) \mathbf{1}_{|y|\le N_2} f(y)\, dy \\ &= \left(\mathbf{1}_{|x|\le N_1} \mathcal{F}^{-1}\left(F(\xi) * \gamma_{1/N}\right) \mathcal{F}\, \mathbf{1}_{|x|\le N_2} f\right)(x)\,,\end{aligned}$$

which is the first part of the assertion.

Since $F_S^* F_S$ acts as convolution with $(d\sigma)^\vee$, we obtain analogously

$$\begin{aligned}\left(\mathbf{1}_{|x|\le N_1} F_S^* F_S \mathbf{1}_{|x|\le N_2} f)\right)(x) &= \mathbf{1}_{|x|\le N_1} \int (d\sigma)^\vee(x-y)\check{\gamma}(|x-y|/N) \mathbf{1}_{|y|\le N_2} f(y)\, dy \\ &= \left(\mathbf{1}_{|x|\le N_1} \mathcal{F}^{-1}\left(d\sigma * \gamma_{1/N}\right) \mathcal{F}\, \mathbf{1}_{|x|\le N_2} f\right)(x)\end{aligned}$$

since

$$\begin{aligned}\left(d\sigma * \gamma_{1/N}\right)^\vee(x-y) &= \int_{\mathbb{R}^d} d\xi\, \mathrm{e}^{2\pi i \xi\cdot(x-y)} \int_S d\sigma(\eta)\, \gamma_{1/N}(\xi-\eta) \\ &= \int_S d\sigma(\eta)\, \mathrm{e}^{2\pi i \eta(x-y)} \check{\gamma}(|x-y|/N) \\ &= (d\sigma)^\vee(x-y)\check{\gamma}(|x-y|/N)\,.\end{aligned}$$

This concludes the proof. □

Appendix E. Hausdorff measures

In the following we summarize some properties of Hausdorff measures following Wolff's notes [200, Chapter 8].

Definition E.1. Fix $\alpha > 0$ and let $E \subseteq \mathbb{R}^d$. For $\varepsilon > 0$, we define

$$H_\alpha^\varepsilon(E) := \inf\{\sum_{j=1}^\infty r_j^\alpha\}$$

where the infimum is taken over all countable coverings of E by balls $B_{x_j}(r_j)$ with $r_j < \varepsilon$.

Clearly, $H_\alpha^\varepsilon(E)$ decreases when $\varepsilon \searrow 0$ and so we define the *(spherical) Hausdorff measure* [71, p. 7]

$$H_\alpha(E) := \lim_{\varepsilon \searrow 0} H_\alpha^\varepsilon(E)\,. \tag{E.1}$$

It is also clear that for $\beta < \alpha$ we have $H_\alpha^\varepsilon(E) \le H_\beta^\varepsilon(E)$ whenever $\varepsilon \le 1$, i.e.,

$$H_\alpha(E) \text{ is a non-increasing function in } \alpha\,. \tag{E.2}$$

Remarks E.2. (1) If $H_\alpha^1(E) = 0$, then $H_\alpha(E) = 0$. This follows from the definition since a covering showing that $H_\alpha^1(E) < \delta$ will necessarily consist of balls of radius $\delta^{1/\alpha}$.

(2) It is also clear that $H_\alpha(E) = 0$ whenever $\alpha > n$ since one can then already cover $\mathbb{R}^d$ by balls $B_{x_j}(r_j)$ such that $\sum_{j=1}^\infty r_j^\alpha$ is arbitrarily small.

Lemma E.3. *Let $E \subseteq \mathbb{R}^d$. Then there is a unique number α_0, called the* Hausdorff dimension *of E or* $\dim(E)$ *such that $H_\alpha(E) = \infty$ if $\alpha < \alpha_0$ and $H_\alpha(E) = 0$ if $\alpha > \alpha_0$.*

Proof. Define α_0 to be the supremum of all α such that $H_\alpha(E) = \infty$. Thus, $H_\alpha(E) = \infty$ if $\alpha < \alpha_0$ by (E.2). Now suppose $\alpha > \alpha_0$, let $\beta \in (\alpha_0, \alpha)$, and define $M := 1 + H_\beta(E) < \infty$. For given $\varepsilon \in (0,1)$, we can therefore find a covering of balls with $\sum_j r_j^\beta \le M$ and $r_j \le \varepsilon$. Thus,

$$\sum_j r_j^\alpha \le \varepsilon^{\alpha-\beta} \sum_j r_j^\beta \le \varepsilon^{\alpha-\beta} M$$

which goes to zero as $\varepsilon \to 0$. Thus, $H_\alpha(E) = 0$ for $\alpha > \alpha_0$. □

Remarks E.4. (1) The set function H_α may seen to be countably additive on Borel sets, i.e., H_α defines a Borel measure. In particular, $H_\alpha(E \cup F) = H_\alpha(E) + H_\alpha(F)$ for compact, disjoint sets E, F. This is part of the reason one considers H_α instead of any other H_α^ε (e.g., H_α^1). We refer to standard references in the area like Carleson's survey [42], Falconer [71], or Mattila [136].

(2) The Borel measure H_d coincides with $|B_0(1)|^{-1}$ times the Lebesgue measure. If $\alpha < d$, then H_α is non-sigma finite. The follows, e.g., by Lemma E.3 which implies that any set of non-zero Lebesgue measure will have infinite H_α measure.

Example E.5. (1) The canonical example is the usual 1/3-Cantor set on $[0,1]$ This has a covering of 2^n intervals of length 3^{-n}, so it has finite $H_{\frac{\log 2}{\log 3}}$ measure. It is not hard to show that in fact its $H_{\frac{\log 2}{\log 3}}$ measure is non-zero. This can be done geometrically (cf. [71, Theorem 8.6][30] with similitudes $\psi_1(x) = x/3$ and $\psi_2(x) = (x+2)/3$[31]), or one can apply Proposition E.7 below to the Cantor measure. In particular, the Hausdorff dimension of the Cantor set is $\log 2/\log 3$.

There are various other notions of dimension. Let us mention only one of them, namely the Minkowski dimension which we define here only for compact sets.

Definition E.6 (Minkowski dimension)**.** Suppose $E \subseteq \mathbb{R}^d$ is compact, then let $E_\delta = \{x \in \mathbb{R}^d : \mathrm{dist}(x, E) < \delta\}$ be the δ-neighborhood of E.

Let α_0 be the supremum of all $\alpha > 0$ such that, for some constant C,

$$|E_\delta| \ge C\delta^{d-\alpha}$$

for all $\delta \in (0,1]$. Then, α_0 is called the *lower Minkowski dimension of* E, denoted by $d_L(E)$.

Let α_1 be the supremum of all $\alpha > 0$ such that, for some constant C,

$$|E_\delta| \ge C\delta^{d-\alpha}$$

for a sequence of δ's converging to zero. Then α_1 is called the *upper Minkowski dimension of* E, denoted by $d_U(E)$.

It would also be possible to define the Minkowski dimensions like the Hausdorff dimension but restricting to coverings of balls *of the same size.* Namely, define a set S to be *δ-separated* if any two distinct points $x, y \in S$ satisfy $|x - y| > \delta$. Let $\mathcal{E}_\delta(E)$ be the *δ-entropy on* E, defined by the maximal possible cardinality for a δ-separated subset of E. [32] Then, one can show that

$$d_L(E) = \liminf_{\delta\to 0} \frac{\log \mathcal{E}_\delta(E)}{\log(1/\delta)},$$

$$d_U(E) = \limsup_{\delta\to 0} \frac{\log \mathcal{E}_\delta(E)}{\log(1/\delta)}.$$

---

[30] See also Corollary 8.7 there which says that the Cantor set is indeed self-similar since it satisfies the open set condition $\bigcup_{j=1}^2 \psi_j([0,1]) \subset [0,1]$.

[31] See also Theorem 8.3 there which says that there is a unique compact set $E \subseteq \mathbb{R}$ such that $\psi(E) := \bigcup_j \psi_j(E) = E$ for any finite set of contractions, and for any non-empty compact set $F \subseteq \mathbb{R}$, one has $\lim_{k\to\infty} \psi^k(F) = E$ in Hausdorff metric.

[32] Show that $\mathcal{E}_\delta(E)$ is comparable to the minimum number of δ-balls required to cover E.

Notice that countable sets may have positive lower Minkowski dimension; consider, e.g., the set $\{1/n\}_{n=1}^{\infty} \cup \{0\}$ which has upper and lower Minkowski dimension 1/2.

In the following, we will give a potential theoretic characterization of the Hausdorff measure. If E is a compact set, then let $P(E)$ denote the space of all probability measures supported on E. The following will be quite helpful.

Proposition E.7 (Frostman's lemma). *Suppose $E \subseteq \mathbb{R}^d$ is compact. Then the following two assertions are equivalent.*

(1) There is a $\mu \in P(E)$ such that

$$\mu(B_x(r)) \leq Cr^{\alpha} \tag{E.3}$$

for a suitable constant C and all $x \in \mathbb{R}^d$, $r > 0$.

(2) $H_\alpha(E) > 0$.

Proof. See Wolff [200, Proposition 8.2] or Mattila [137, Theorem 2.7]. □

Let us now define the *α-dimensional energy* of a (positive) measure μ with compact support [33] by the formula

$$I_\alpha(\mu) := \int \frac{d\mu(x)d\mu(y)}{|x-y|^{\alpha}}, \quad \alpha < d. \tag{E.4}$$

Let us also define the mean field potential

$$V_\mu^\alpha(x) := \int |x-y|^{-\alpha} d\mu(y),$$

i.e., we have

$$I_\alpha(\mu) = \int V_\mu^\alpha(x)\, d\mu(x). \tag{E.5}$$

Roughly, one expects μ to have $I_\alpha(\mu) < \infty$ if and only if it satisfies (E.3). Although this precise statement is false, we will now see that the Hausdorff dimension of a compact subset can still be defined in terms of the energies of measures in $P(E)$.

Lemma E.8. *Let μ be a probability measure with compact support. Then, the following two assertions hold.*

(1) If μ satisfies (E.3)*, then $I_\beta(\mu) < \infty$ for all $\beta < \alpha$.*

(2) Conversely, if μ satisfies $I_\alpha(\mu) < \infty$, then there is another probability measure ν such that $\nu(X) \leq 2\mu(X)$ for all sets X and such that ν satisfies (E.3)*.*

Proof. (1) Without loss of generality, we assume that the diameter of $\mathrm{supp}(\mu)$ is ≤ 1. Then, by (E.3),

$$\int V_\mu^\beta(x)\, d\mu(x) \lesssim \int \sum_{j=0}^{\infty} 2^{j\beta} \mu(B_x(2^{-j}))\, d\mu(x) \lesssim \int \sum_{j=0}^{\infty} 2^{j(\beta-\alpha)}\, d\mu(x) \lesssim 1.$$

(2) Let $F = \{x : V_\mu^\alpha(x) \leq 2I_\alpha(\mu)\}$, then $\mu(F) \geq 1/2$ by $I_\alpha(\mu) = \int V_\mu^\alpha(x)\, d\mu(x)$ (and the mean value theorem). Let us now define the new probability measure ν by $\nu(X) = \mu(X \cap F)/\mu(F)$. By the previous argument $\nu(X) \leq 2\mu(X)$ and we are left to show that ν satisfies (E.3). Suppose first $x \in F$. If $r > 0$ then

$$r^{-\alpha}\nu(B_x(r)) \leq V_\nu^\alpha(x) \leq 2V_\mu^\alpha(x) \leq 4I_\alpha(\mu)$$

[33] The compact support assumption is not needed; it is only included to simplify the presentation

which shows (E.3) whenever $x \in F$. For general x we distinguish between the cases where the intersection $B_x(r) \cap F$ is empty or not. Assume first that r is such that $B_x(r) \cap F = \emptyset$. Then evidently $\nu(B_x(r)) = 0$. Else, if $B_x(r) \cap F \neq \emptyset$, let $y \in B_x(r) \cap F$ and observe that $\nu(B_x(r)) \leq \nu(B_y(2r)) \lesssim r^\alpha$ by the first part of the proof. □

We are now ready to give an alternative characterization of Hausdorff dimension for compact subsets of $\mathbb{R}^d$.

Proposition E.9. *If E is compact then the Hausdorff dimension of E coincides with the number*

$$\sup\{\alpha : \exists \mu \in P(E) \text{ with } I_\alpha(\mu) < \infty\} . \tag{E.6}$$

Proof. Denote the above supremum by s. If $\beta < s$, then by (2) of the previous lemma, we know that E supports a measure with $\mu(B_x(r)) \leq Cr^\beta$. But then by Proposition E.7, we have $H_\beta(E) > 0$, i.e., $\beta \leq \dim E$ which means $s \leq \dim E$. Conversely, if $\beta < \dim E$, then by Proposition E.7, E supports a measure with $\mu(B_x(r)) \leq Cr^{\beta+\varepsilon}$ for some sufficiently small $\varepsilon > 0$. Then $I_\beta(\mu) < \infty$ and so $\beta \leq s$ which shows $\dim E \leq s$. □

As the α-energy is the expectation value of a translational invariant function, the Fourier transform should come in handy. In particular, we will make us of the elementary

Proposition E.10. *Let μ be a positive measure with compact support and $\alpha < d$. Then*

$$I_\alpha(\mu) = \int \frac{d\mu(x)d\mu(y)}{|x-y|^\alpha} = c_\alpha \int_{\mathbb{R}^d} \frac{|\hat{\mu}(\xi)|^2}{|\xi|^{d-\alpha}}\, d\xi\,, \quad \text{where } c_\alpha = \frac{\Gamma\left(\frac{d-\alpha}{2}\right)\pi^{\alpha-d/2}}{\Gamma(\alpha/2)}\,. \tag{E.7}$$

Using this and Proposition E.9 allows us to prove a lower bound on the Hausdorff dimension of the support of probability measures.

Corollary E.11. *Suppose μ is a compactly supported probability measure on $\mathbb{R}^d$ with*

$$|\hat{\mu}(\xi)| \lesssim |\xi|^{-\beta} \tag{E.8}$$

for some $0 < \beta < d/2$, or more generally that (E.8) *is true in the L^2 sense*

$$\int_{B_0(N)} |\hat{\mu}(\xi)|^2\, d\xi \lesssim N^{d-2\beta}\,. \tag{E.9}$$

Then the dimension of the support of μ is at least 2β.

Proof. By Proposition E.9 it suffices to show that if (E.9) holds, then $I_\alpha(\mu) < \infty$ for all $\alpha < 2\beta$. But in view of the Fourier representation of $I_\alpha(\mu)$, we have (using $|\hat{\mu}(\xi)| \leq \|\mu\|_1 = 1$)

$$\begin{aligned} c_\alpha^{-1} I_\alpha(\mu) &= \left(\int_{|\xi|\leq 1} + \int_{|\xi|\geq 1}\right) \frac{|\hat{\mu}(\xi)|^2}{|\xi|^{d-\alpha}}\, d\xi \lesssim \sup_\xi |\hat{\mu}(\xi)|^2 + \sum_{j=0}^\infty 2^{-j(d-\alpha)} \int\limits_{2^j\leq|\xi|\leq 2^{j+1}} |\hat{\mu}(\xi)|^2\, d\xi \\ &\lesssim \|\mu\|_1^2 + \sum_{j=0}^\infty 2^{-j(d-\alpha)+j(d-2\beta)} < \infty\,, \end{aligned}$$

whenever $\alpha < 2\beta$. This shows (E.9) and concludes the proof. □

One may ask the converse question, whether a compact set with dimension α *must* support a measure μ satisfying

$$|\hat{\mu}(\xi)| \lesssim_\varepsilon (1+|\xi|)^{-\alpha/2-\varepsilon} \tag{E.10}$$

for all $\varepsilon > 0$. The answer is (emphatically) no [34]. Indeed, there are many counterexamples, i.e., sets with positive Hausdorff dimension which do not support any measure whose Fourier transform even decays as $|\xi| \to \infty$. Consider, e.g., the line segment $E = [0,1] \times \{0\} \subseteq \mathbb{R}^2$. Clearly, E has dimension 1, but if μ is a measure supported on E, then $\hat{\mu}(\xi)$ only on ξ_1, and so it cannot go to zero as $\xi_1^2 + \xi_2^2 \to 0$. If one considers only the case $d = 1$, this question is related to the classical question of "sets of uniqueness", see, e.g., Salem [151] or Zygmund [206]. For instance, one can show that the standard 1/3 Cantor set does not support any measure such that $\hat{\mu}$ vanishes at infinity. Indeed, it is non-trivial to show that a "non-counterexample" exists, i.e., a set E with given dimension α which supports a measure satisfying (E.10). One can find a construction of such a set due to Kaufman in Wolff [200, Chapter 9].

Remark E.12. There is an important relation between the Fourier transform of Borel measures and dynamical properties thereof in quantum mechanics. Consider a self-adjoint Hamiltonian H in some Hilbert space $\mathcal{H}$, the associated spectral measure (on Borel sets in $\mathbb{R}$) $d\mu_\psi(\lambda) = (\psi, dE_H(\lambda)\psi)$ for some $\psi \in \mathcal{H}$, and its Fourier transform

$$\hat{\mu}_\psi(t) = \int_{\mathbb{R}} \mathrm{e}^{it\lambda}\, d\mu_\psi = (\psi, \mathrm{e}^{itH}\psi) = (\psi(0), \psi(t))\,.$$

Its absolute square, i.e., $|\hat{\mu}_\psi(t)| = |(\psi(0), \psi(t))|^2$, denotes the *survival probability* as $\psi(0)$ is evolved along the Hamiltonian flow. Usually, one is interested in its Cesaro average

$$< |\hat{\mu}|^2 >_T := \frac{1}{T}\int_0^T |\hat{\mu}(t)|^2\, dt\,.$$

Wiener's theorem then asserts $\lim_{T\to\infty} < |\hat{\mu}_\psi|^2 >_T = \sum_{\lambda\in\mathbb{R}} |\mu_\psi(\{\lambda\})|^2$. Thus, if $\psi \in \mathcal{H}_c$, the continuous spectral subspace of H, the survival probability decays to zero. If μ_ψ is uniformly α Hölder continuous (UαH), i.e., $\mu_\psi(I) \leq |I|^\alpha$ for some $\alpha \in [0,1]$ and where $|I|$ denotes the Lebesgue measure, then, Strichartz's theorem [173] (see also Last [128, Theorem 3.1]) refines Wiener's theorem and says

$$< |\hat{\mu}_\psi|^2 >_T \lesssim T^{-\alpha}\,, \quad \psi \in \mathcal{H}_{uh}(\alpha)$$

where $\mathcal{H}_{uh}(\alpha) = \{\psi : \mu_\psi \text{ is U}\alpha\text{H}\}$.

We say that a measure μ *is α-continuous* iff $\mu(E) = 0$ for any set E for which the Hausdorff measure $H_\alpha(E) = 0$ and denote $\mathcal{H}_{\alpha c} = \{\psi : \mu_\psi \text{ is } \alpha\text{-continuous}\}$. Last [128, Theorem 5.2] showed that for all $\alpha \in [0,1]$ one has $\overline{\mathcal{H}_{uh}(\alpha)} = \mathcal{H}_{\alpha c}$ which means that $\mathcal{H}_{\alpha c}$ must have a dense subset of vectors for which $\sup_T T^\alpha < |\hat{\mu}_\psi|^2 >_T < \infty$.

Moreover, the α-dimensional energy defined in (E.4) is related, via the Fourier transform to

$$\int_0^\infty \frac{|\hat{\mu}_\psi(t)|^2}{t^{1-\alpha}}\, dt = (\hat{\mu}_\psi, |\cdot|^{-1+\alpha}\hat{\mu}_\psi) = (d\mu_\psi, |\cdot|^{-\alpha} * d\mu_\psi) = \int \frac{d\mu_\psi(x)d\mu_\psi(y)}{|x-y|^\alpha} = I_\alpha(\mu_\psi)\,.$$

Recalling Proposition E.9 it is then interesting to observe that (cf. Last [128, Lemma 5.1] μ_ψ is α-continuous, whenever $I_\alpha(\mu_\psi) < \infty$.

References

[1] Robert A. Adams and John J. F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] Shiri Artstein-Avidan, Apostolos Giannopoulos, and Vitali D. Milman. *Asymptotic Geometric Analysis. Part I*, volume 202 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[34] On the other hand, if one interprets decay in an L^2 averaged sense, the answer becomes 'yes' since the calculation on the proof of the above corollary is clearly reversible.

[3] Jong-Guk Bak and Andreas Seeger. Extensions of the Stein-Tomas theorem. *Math. Res. Lett.*, 18(4):767–781, 2011.

[4] William Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[5] Colin Bennett and Robert Sharpley. *Interpolation of Operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1988.

[6] Jonathan Bennett. Aspects of multilinear harmonic analysis related to transversality. In *Harmonic Analysis and Partial Differential Equations*, volume 612 of *Contemp. Math.*, pages 1–28. Amer. Math. Soc., Providence, RI, 2014.

[7] Jonathan Bennett, Anthony Carbery, and Terence Tao. On the multilinear restriction and Kakeya conjectures. *Acta Math.*, 196(2):261–302, 2006.

[8] Jöran Bergh and Jörgen Löfström. *Interpolation Spaces. An Introduction*. Springer-Verlag, Berlin-New York, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[9] A. S. Besicovitch. On Kakeya's problem and a similar one. *Math. Z.*, 27(1):312–320, 1928.

[10] Rajendra Bhatia. *Matrix Analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.

[11] M. Sh. Birman and D. R. Yafaev. Asymptotic behavior of the spectrum of the scattering matrix. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 110:3–29, 241, 1981. Boundary value problems of mathematical physics and related questions in the theory of functions, 13.

[12] M. Sh. Birman and D. R. Yafaev. Asymptotic behavior of the spectrum of the scattering matrix. *Journal of Soviet Mathematics*, 25(1):793–814, 1984.

[13] Sabine Bögli and Jean-Claude Cuenin. Counterexample to the Laptev–Safronov conjecture. *arXiv e-prints*, page arXiv:2109.06135, September 2021.

[14] Jean-Marc Bouclet and Haruya Mizutani. Uniform resolvent and Strichartz estimates for Schrödinger equations with critical singularities. *Trans. Amer. Math. Soc.*, 370(10):7293–7333, 2018.

[15] J. Bourgain. Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.*, 1(2):147–187, 1991.

[16] J. Bourgain. L^p-estimates for oscillatory integrals in several variables. *Geom. Funct. Anal.*, 1(4):321–374, 1991.

[17] J. Bourgain. A remark on Schrödinger operators. *Israel J. Math.*, 77(1-2):1–16, 1992.

[18] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.*, 3(2):107–156, 1993.

[19] J. Bourgain. On the distribution of Dirichlet sums. *J. Anal. Math.*, 60:21–32, 1993.

[20] J. Bourgain. Random lattice Schrödinger operators with decaying potential: some higher dimensional phenomena. In *Geometric Aspects of Functional Analysis*, volume 1807 of *Lecture Notes in Math.*, pages 70–98. Springer, Berlin, 2003.

[21] J. Bourgain. On Strichartz's inequalities and the nonlinear Schrödinger equation on irrational tori. In *Mathematical Aspects of Nonlinear Dispersive Equations*, volume 163 of *Ann. of Math. Stud.*, pages 1–20. Princeton Univ. Press, Princeton, NJ, 2007.

[22] J. Bourgain. Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces. *Israel J. Math.*, 193(1):441–458, 2013.

[23] J. Bourgain. A note on the Schrödinger maximal function. *J. Anal. Math.*, 130:393–396, 2016.

[24] J. Bourgain, J. Lindenstrauss, and V. Milman. Approximation of zonoids by zonotopes. *Acta Math.*, 162(1-2):73–141, 1989.

[25] Jean Bourgain. Estimations de certaines fonctions maximales. *C. R. Acad. Sci. Paris Sér. I Math.*, 301(10):499–502, 1985.

[26] Jean Bourgain. On the restriction and multiplier problems in $\mathbf{R}^3$. In *Geometric Aspects of Functional Analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 179–191. Springer, Berlin, 1991.

[27] Jean Bourgain. Some new estimates on oscillatory integrals. In *Essays on Fourier Analysis in Honor of Elias M. Stein (Princeton, NJ, 1991)*, volume 42 of *Princeton Math. Ser.*, pages 83–112. Princeton Univ. Press, Princeton, NJ, 1995.

[28] Jean Bourgain. On random Schrödinger operators on $\mathbb{Z}^2$. *Discrete Contin. Dyn. Syst.*, 8(1):1–15, 2002.

[29] Jean Bourgain and Ciprian Demeter. The proof of the ℓ^2 decoupling conjecture. *Ann. of Math. (2)*, 182(1):351–389, 2015.

[30] Jean Bourgain and Ciprian Demeter. A study guide for the ℓ^2 decoupling theorem. *Chin. Ann. Math. Ser. B*, 38(1):173–200, 2017.

[31] Jean Bourgain and Larry Guth. Bounds on oscillatory integral operators based on multilinear estimates. *Geom. Funct. Anal.*, 21(6):1239–1295, 2011.

[32] Nicolas Burq, Fabrice Planchon, John G. Stalker, and A. Shadi Tahvildar-Zadeh. Strichartz estimates for the wave and Schrödinger equations with the inverse-square potential. *J. Funct. Anal.*, 203(2):519–549, 2003.

[33] Nicolas Burq, Fabrice Planchon, John G. Stalker, and A. Shadi Tahvildar-Zadeh. Strichartz estimates for the wave and Schrödinger equations with potentials of critical decay. *Indiana Univ. Math. J.*, 53(6):1665–1680, 2004.

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

[35] Anthony Carbery. The boundedness of the maximal Bochner-Riesz operator on $L^4(\mathbf{R}^2)$. *Duke Math. J.*, 50(2):409–416, 1983.

[36] Anthony Carbery. Restriction implies Bochner-Riesz for paraboloids. *Math. Proc. Cambridge Philos. Soc.*, 111(3):525–529, 1992.

[37] Anthony Carbery. Almost-orthogonality in the Schatten-von Neumann classes. *J. Operator Theory*, 62(1):151–158, 2009.

[38] Anthony Carbery. A remark on reverse Littlewood-Paley, restriction and Kakeya. *arXiv e-prints*, page arXiv:1507.02515, July 2015.

[39] Anthony Carbery, José L. Rubio de Francia, and Luis Vega. Almost everywhere summability of Fourier integrals. *J. London Math. Soc. (2)*, 38(3):513–524, 1988.

[40] Anthony Carbery, Andreas Seeger, Stephen Wainger, and James Wright. Classes of singular integral operators along variable lines. *J. Geom. Anal.*, 9(4):583–605, 1999.

[41] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.

[42] Lennart Carleson. *Selected Problems on Exceptional Sets.* Van Nostrand Mathematical Studies, No. 13. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1967.

[43] Lennart Carleson. Some analytic problems related to statistical mechanics. In *Euclidean Harmonic Analysis (Proc. Sem., Univ. Maryland, College Park, Md., 1979)*, volume 779 of *Lecture Notes in Math.*, pages 5–45. Springer, Berlin, 1980.

[44] Lennart Carleson and Per Sjölin. Oscillatory integrals and a multiplier problem for the disc. *Studia Math.*, 44:287–299. (errata insert), 1972.

[45] Peng Chen, El Maati Ouhabaz, Adam Sikora, and Lixin Yan. Restriction estimates, sharp spectral multipliers and endpoint estimates for Bochner-Riesz means. *J. Anal. Math.*, 129:219–283, 2016.

[46] Peng Chen, El Maati Ouhabaz, Adam Sikora, and Lixin Yan. Spectral multipliers without semigroup framework and application to random walks. *J. Math. Pures Appl. (9)*, 143:162–191, 2020.

[47] Michael Christ and Alexander Kiselev. Maximal functions associated to filtrations. *J. Funct. Anal.*, 179(2):409–425, 2001.

[48] E. Compaan, R. Lucá, and G. Staffilani. Pointwise convergence of the Schrödinger flow, 2019.

[49] A. Córdoba. Multipliers of $\mathcal{F}(L^p)$. In *Euclidean Harmonic Analysis (Proc. Sem., Univ. Maryland, College Park, Md., 1979)*, volume 779 of *Lecture Notes in Math.*, pages 162–177. Springer, Berlin, 1980.

[50] A. Cordoba and R. Fefferman. On the equivalence between the boundedness of certain classes of maximal and multiplier operators in Fourier analysis. *Proc. Nat. Acad. Sci. U.S.A.*, 74(2):423–425, 1977.

[51] Antonio Cordoba. The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.*, 99(1):1–22, 1977.

[52] Antonio Córdoba. Geometric Fourier analysis. *Ann. Inst. Fourier (Grenoble)*, 32(3):vii, 215–226, 1982.

[53] Jean-Claude Cuenin. Eigenvalue bounds for Dirac and fractional Schrödinger operators with complex potentials. *J. Funct. Anal.*, 272(7):2987–3018, 2017.

[54] Jean-Claude Cuenin. From spectral cluster to uniform resolvent estimates on compact manifolds. *arXiv e-prints*, page arXiv:2011.07254, November 2020.

[55] Jean-Claude Cuenin. Effective upper bounds on the number of resonance in potential scattering. *arXiv e-prints*, page arXiv:2209.06079, September 2022.

[56] Jean-Claude Cuenin and Konstantin Merz. Weak coupling limit for Schrödinger-type operators with degenerate kinetic energy for a large class of potentials. *arXiv e-prints*, page arXiv:2006.07110, June 2020.

[57] Jean-Claude Cuenin and Konstantin Merz. On the number and sums of eigenvalues of Schrödinger-type operators with degenerate kinetic energy. *arXiv e-prints. To appear in Operator Theory: Advances and Applications*, page arXiv:2202.05212, February 2022.

[58] Jean-Claude Cuenin and Konstantin Merz. Random Schrödinger operators with complex decaying potentials. *arXiv e-prints*, page arXiv:2201.04466, January 2022.

[59] F. Cunningham, Jr. The Kakeya problem for simply connected and for star-shaped sets. *Amer. Math. Monthly*, 78:114–129, 1971.

[60] Björn E. J. Dahlberg and Carlos E. Kenig. A note on the almost everywhere behavior of solutions to the Schrödinger equation. In *Harmonic Analysis (Minneapolis, Minn., 1981)*, volume 908 of *Lecture Notes in Math.*, pages 205–209. Springer, Berlin-New York, 1982.

[61] Miguel de Guzmán. *Real Variable Methods in Fourier Analysis*, volume 46 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York, 1981. Notas de Matemática [Mathematical Notes], 75.

[62] Ciprian Demeter. Incidence theory and restriction estimates. *arXiv e-prints*, page arXiv:1401.1873, January 2014.

[63] Ciprian Demeter. *Fourier Restriction, Decoupling, and Applications*, volume 184 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2020.

[64] Evangelos Dimou and Andreas Seeger. On pointwise convergence of Schrödinger means, 2019.

[65] Xiumin Du, Larry Guth, and Xiaochun Li. A sharp Schrödinger maximal estimate in $\mathbb{R}^2$. *Ann. of Math. (2)*, 186(2):607–640, 2017.

[66] Xiumin Du, Larry Guth, Xiaochun Li, and Ruixiang Zhang. Pointwise convergence of Schrödinger solutions and multilinear refined Strichartz estimates. *Forum Math. Sigma*, 6:e14, 18, 2018.

[67] Xiumin Du and Ruixiang Zhang. Sharp L^2 estimates of the Schrödinger maximal function in higher dimensions. *Ann. of Math. (2)*, 189(3):837–861, 2019.

[68] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[69] Zeev Dvir. On the size of Kakeya sets in finite fields. *J. Amer. Math. Soc.*, 22(4):1093–1097, 2009.

[70] Semyon Dyatlov and Maciej Zworski. *Mathematical Theory of Scattering Resonances*, volume 200 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2019.

[71] K. J. Falconer. *The Geometry of Fractal Sets*, volume 85 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1986.

[72] Charles Fefferman. Inequalities for strongly singular convolution operators. *Acta Math.*, 124:9–36, 1970.

[73] Charles Fefferman. The multiplier problem for the ball. *Ann. of Math. (2)*, 94:330–336, 1971.

[74] Charles Fefferman. A note on spherical summation multipliers. *Israel J. Math.*, 15:44–52, 1973.

[75] Charles Fefferman. Selected theorems by Eli Stein. In *Essays on Fourier Analysis in Honor of Elias M. Stein (Princeton, NJ, 1991)*, volume 42 of *Princeton Math. Ser.*, pages 1–35. Princeton Univ. Press, Princeton, NJ, 1995.

[76] Charles L. Fefferman. The uncertainty principle. *Bull. Amer. Math. Soc. (N.S.)*, 9(2):129–206, 1983.

[77] T. Figiel, J. Lindenstrauss, and V. D. Milman. The dimension of almost spherical sections of convex bodies. *Acta Math.*, 139(1-2):53–94, 1977.

[78] Gerald B. Folland. *Harmonic Analysis in Phase Space*, volume 122 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989.

[79] Gerald B. Folland. *Real Analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern Techniques and Their Applications, A Wiley-Interscience Publication.

[80] Gerald B. Folland and Alladi Sitaram. The uncertainty principle: a mathematical survey. *J. Fourier Anal. Appl.*, 3(3):207–238, 1997.

[81] D. Foschi and D. Oliveira e Silva. Some recent progress on sharp Fourier restriction theory. *Anal. Math.*, 43(2):241–265, 2017.

[82] Damiano Foschi and Sergiu Klainerman. Bilinear space-time estimates for homogeneous wave equations. *Ann. Sci. École Norm. Sup. (4)*, 33(2):211–274, 2000.

[83] Rupert L. Frank. Eigenvalue bounds for Schrödinger operators with complex potentials. *Bull. Lond. Math. Soc.*, 43(4):745–750, 2011.

[84] Rupert L. Frank. Eigenvalue bounds for Schrödinger operators with complex potentials. III. *Trans. Amer. Math. Soc.*, 370(1):219–240, 2018.

[85] Rupert L. Frank, Ari Laptev, Elliott H. Lieb, and Robert Seiringer. Lieb-Thirring inequalities for Schrödinger operators with complex-valued potentials. *Lett. Math. Phys.*, 77(3):309–316, 2006.

[86] Rupert L. Frank, Ari Laptev, and Oleg Safronov. On the number of eigenvalues of Schrödinger operators with complex potentials. *J. Lond. Math. Soc. (2)*, 94(2):377–390, 2016.

[87] Rupert L. Frank and Julien Sabin. Restriction theorems for orthonormal functions, Strichartz inequalities, and uniform Sobolev estimates. *Amer. J. Math.*, 139(6):1649–1691, 2017.

[88] Rupert L. Frank and Barry Simon. Eigenvalue bounds for Schrödinger operators with complex potentials. II. *J. Spectr. Theory*, 7(3):633–658, 2017.

[89] K. O. Friedrichs. *Pseudo-Differential Operators. An introduction.* Notes prepared with the assistance of R. Vaillancourt. Revised edition. Courant Institute of Mathematical Sciences, New York University, New York, 1970.

[90] Emilio Gagliardo. Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili. *Rend. Sem. Mat. Univ. Padova*, 27:284–305, 1957.

[91] Gustavo Garrigós and Andreas Seeger. A mixed norm variant of Wolff's inequality for paraboloids. In *Harmonic Analysis and Partial Differential Equations*, volume 505 of *Contemp. Math.*, pages 179–197. Amer. Math. Soc., Providence, RI, 2010.

[92] I. M. Gel′fand and G. E. Shilov. *Generalized Functions. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 2016. Properties and Operations, Translated from the 1958 Russian original by Eugene Saletan, Reprint of the 1964 English translation.

[93] F. Gesztesy, Y. Latushkin, M. Mitrea, and M. Zinchenko. Nonselfadjoint operators, infinite determinants, and some applications. *Russ. J. Math. Phys.*, 12(4):443–471, 2005.

[94] F. Gesztesy, Y. Latushkin, M. Mitrea, and M. Zinchenko. Erratum to "Non self adjoint operators, infinite determinants, and some applications," Russ. J. Math. Phys. 12, 443–471 (2005). *Russ. J. Math. Phys.*, 27(3):410, 2020.

[95] J. Ginibre and G. Velo. Smoothing properties and retarded estimates for some dispersive evolution equations. *Comm. Math. Phys.*, 144(1):163–188, 1992.

[96] Loukas Grafakos. *Classical Fourier Analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[97] Loukas Grafakos. *Modern Fourier Analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[98] Philip T Gressman, Shaoming Guo, Lillian B Pierce, Joris Roos, and Po-Lam Yung. Reversing a philosophy: from counting to square functions and decoupling. *The Journal of Geometric Analysis*, pages 1–21, 2021. In press.

[99] Colin Guillarmou, Andrew Hassell, and Adam Sikora. Restriction and spectral multiplier theorems on asymptotically conic manifolds. *Anal. PDE*, 6(4):893–950, 2013.

[100] Larry Guth. The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture. *Acta Math.*, 205(2):263–286, 2010.

[101] Larry Guth. A short proof of the multilinear Kakeya inequality. *Math. Proc. Cambridge Philos. Soc.*, 158(1):147–153, 2015.

[102] Larry Guth. *Polynomial Methods in Combinatorics*, volume 64 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2016.

[103] Larry Guth. A restriction estimate using polynomial partitioning. *J. Amer. Math. Soc.*, 29(2):371–413, 2016.

[104] Victor Havin and Burglind Jöricke. *The Uncertainty Principle in Harmonic Analysis*, volume 28 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1994.

[105] C. S. Herz. Fourier transforms related to convex sets. *Ann. of Math. (2)*, 75:81–92, 1962.

[106] Carl S. Herz. On the mean inversion of Fourier and Hankel transforms. *Proc. Nat. Acad. Sci. U.S.A.*, 40:996–999, 1954.

[107] Jonathan Hickman and Marco Vitturi. Notes on the solution of the ℓ^2 decoupling conjecture by Bourgain and Demeter. Available at `https://www.math.sciences.univ-nantes.fr/~vitturi/lecture_notes/lecture_notes.html`, 2015.

[108] Lars Hörmander. Estimates for translation invariant operators in L^p spaces. *Acta Math.*, 104:93–140, 1960.

[109] Lars Hörmander. *The Analysis of Linear Partial Differential Operators. II*, volume 257 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1983. Differential Operators with Constant Coefficients.

[110] Lars Hörmander. *The Analysis of Linear Partial Differential Operators. III*, volume 274 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985. Pseudodifferential Operators.

[111] Lars Hörmander. *The Analysis of Linear Partial Differential Operators. I.* Springer Study Edition. Springer-Verlag, Berlin, second edition, 1990. Distribution Theory and Fourier analysis.

[112] Shanlin Huang, Xiaohua Yao, and Quan Zheng. Remarks on L^p-limiting absorption principle of Schrödinger operators and applications to spectral multiplier theorems. *Forum Math.*, 30(1):43–55, 2018.

[113] Teruo Ikebe. Eigenfunction expansions associated with the Schroedinger operators and their applications to scattering theory. *Arch. Rational Mech. Anal.*, 5:1–34 (1960), 1960.

[114] Alexandru D. Ionescu and Wilhelm Schlag. Agmon-Kato-Kuroda theorems for a large class of perturbations. *Duke Math. J.*, 131(3):397–440, 2006.

[115] J.-P. Kahane. Trois notes sur les ensembles parfaits linéaires. *Enseign. Math. (2)*, 15:185–192, 1969.

[116] Tosio Kato. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1 edition, 1966.

[117] Nets Katz and Terence Tao. Recent progress on the Kakeya conjecture. In *Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000)*, number Vol. Extra, pages 161–179, 2002.

[118] Nets Hawk Katz and Terence Tao. New bounds for Kakeya problems. *J. Anal. Math.*, 87:231–263, 2002. Dedicated to the memory of Thomas H. Wolff.

[119] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

[120] C. E. Kenig, A. Ruiz, and C. D. Sogge. Uniform Sobolev inequalities and unique continuation for second order constant coefficient differential operators. *Duke Math. J.*, 55(2):329–347, 1987.

[121] Carlos E. Kenig and Alberto Ruiz. A strong type (2, 2) estimate for a maximal operator associated to the Schrödinger equation. *Trans. Amer. Math. Soc.*, 280(1):239–246, 1983.

[122] Seongyeon Kim, Ihyeok Seo, and Jihyeon Seok. On local-in-time Strichartz estimates for the Schrödinger equation with singular potentials. *Math. Nachr.*, 294(9):1765–1772, 2021.

[123] S. Klainerman and M. Machedon. Space-time estimates for null forms and the local existence theorem. *Comm. Pure Appl. Math.*, 46(9):1221–1268, 1993.

[124] Herbert Koch and Daniel Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.

[125] Herbert Koch and Daniel Tataru. Carleman estimates and absence of embedded eigenvalues. *Comm. Math. Phys.*, 267(2):419–449, 2006.

[126] Steven G. Krantz. *A Panorama of Harmonic Analysis*, volume 27 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC, 1999.

[127] Ari Laptev and Oleg Safronov. Eigenvalue estimates for Schrödinger operators with complex potentials. *Comm. Math. Phys.*, 292(1):29–54, 2009.

[128] Yoram Last. Quantum dynamics and decompositions of singular continuous spectra. *J. Funct. Anal.*, 142(2):406–445, 1996.

[129] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and Processes.

[130] Xiaochun Li and Shukun Wu. New estimates of maximal Bochner-Riesz operator in the plane, 2019.

[131] Elliott H. Lieb. Gaussian kernels have only Gaussian maximizers. *Invent. Math.*, 102(1):179–208, 1990.

[132] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[133] Walter Littman. Fourier transforms of surface-carried measures and differentiability of surface averages. *Bull. Amer. Math. Soc.*, 69:766–770, 1963.

[134] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc*, 55:961–962, 1949.

[135] André Martinez. *An Introduction to Semiclassical and Microlocal Analysis*. Universitext. Springer-Verlag, New York, 2002.

[136] Pertti Mattila. *Geometry of Sets and Measures in Euclidean Spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[137] Pertti Mattila. *Fourier Analysis and Hausdorff Dimension*, volume 150 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015.

[138] Vitali D. Milman and Gideon Schechtman. *Asymptotic Theory of Finite-Dimensional Normed Spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[139] Haruya Mizutani and Xiaohua Yao. Kato smoothing, Strichartz and uniform Sobolev estimates for fractional operators with sharp Hardy potentials. *Comm. Math. Phys.*, 388(1):581–623, 2021.

[140] A. Moyua and L. Vega. Bounds for the maximal function associated to periodic solutions of one-dimensional dispersive equations. *Bull. Lond. Math. Soc.*, 40(1):117–128, 2008.

[141] Camil Muscalu and Wilhelm Schlag. *Classical and Multilinear Harmonic Analysis. Vol. I*, volume 137 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.

[142] F. W. J. Olver. Bessel functions of integer order. In Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, chapter 9, pages 355–433. Dover Publications, New York, 5 edition, 1968.

[143] Francesco Orabona and David Pal. Optimal non-asymptotic lower bound on the minimax regret of learning with expert advice. *arXiv e-prints*, page arXiv:1511.02176, November 2015.

[144] Alain Pajor and Nicole Tomczak-Jaegermann. Subspaces of small codimension of finite-dimensional Banach spaces. *Proc. Amer. Math. Soc.*, 97(4):637–642, 1986.

[145] Oskar Perron. Über einen Satz von Besicovitsch. *Math. Z.*, 28(1):383–386, 1928.

[146] D. H. Phong and E. M. Stein. Hilbert integrals, singular integrals, and Radon transforms. I. *Acta Math.*, 157(1-2):99–157, 1986.

[147] D. H. Phong and E. M. Stein. Hilbert integrals, singular integrals, and Radon transforms. II. *Invent. Math.*, 86(1):75–113, 1986.

[148] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[149] Alberto Ruiz. Lecture notes: Harmonic analysis and inverse problems. Available at `https://www.researchgate.net/profile/Alberto-Ruiz-21/publication/267937045_HARMONIC_ANALYSIS_AND_INVERSE_PROBLEMS/links/5492f8be0cf22d7925d5876a/HARMONIC-ANALYSIS-AND-INVERSE-PROBLEMS.pdf`, 2002.

[150] Oleg Safronov. Estimates for eigenvalues of the Schrödinger operator with a complex potential. *Bull. Lond. Math. Soc.*, 42(3):452–456, 2010.

[151] Raphaël Salem. *Algebraic Numbers and Fourier Analysis*. D. C. Heath and Co., Boston, Mass., 1963.

[152] W. Schlag, C. Shubin, and T. Wolff. Frequency concentration and location lengths for the Anderson model at small disorders. volume 88, pages 173–220. 2002. Dedicated to the memory of Tom Wolff.

[153] C. Shubin, R. Vakilian, and T. Wolff. Some harmonic analysis questions suggested by Anderson-Bernoulli models. *Geom. Funct. Anal.*, 8(5):932–964, 1998.

[154] Adam Sikora, Lixin Yan, and Xiaohua Yao. Sharp spectral multipliers for operators satisfying generalized Gaussian estimates. *J. Funct. Anal.*, 266(1):368–409, 2014.

[155] Adam Sikora, Lixin Yan, and Xiaohua Yao. Spectral multipliers, Bochner-Riesz means and uniform Sobolev inequalities for elliptic operators. *Int. Math. Res. Not. IMRN*, (10):3070–3121, 2018.

[156] Barry Simon. *Trace Ideals and Their Applications*, volume 120 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, second edition, 2005.

[157] Per Sjölin. Regularity of solutions to the Schrödinger equation. *Duke Math. J.*, 55(3):699–715, 1987.

[158] Hart F. Smith and Christopher D. Sogge. Global Strichartz estimates for nontrapping perturbations of the Laplacian. *Comm. Partial Differential Equations*, 25(11-12):2171–2183, 2000.

[159] Christopher D. Sogge. *Lectures on Non-Linear Wave Equations*. International Press, Boston, MA, second edition, 2008.

[160] Christopher D. Sogge. *Hangzhou Lectures on Eigenfunctions of the Laplacian*, volume 188 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2014.

[161] Christopher D. Sogge. *Fourier Integrals in Classical Analysis*, volume 210 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 2017.

[162] E. M. Stein. On limits of sequences of operators. *Annals of Mathematics*, 74(1):140–170, 1961.

[163] E. M. Stein. Some problems in harmonic analysis. In *Harmonic Analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1*, Proc. Sympos. Pure Math., XXXV, Part, pages 3–20. Amer. Math. Soc., Providence, R.I., 1979.

[164] E. M. Stein. Oscillatory integrals in Fourier analysis. In *Beijing Lectures in Harmonic Analysis (Beijing, 1984)*, volume 112 of *Ann. of Math. Stud.*, pages 307–355. Princeton Univ. Press, Princeton, NJ, 1986.

[165] Elias M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[166] Elias M. Stein. *Topics in Harmonic Analysis Related to the Littlewood-Paley Theory*. Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970.

[167] Elias M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[168] Elias M. Stein and Rami Shakarchi. *Functional Analysis*, volume 4 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2011. Introduction to further topics in analysis.

[169] Elias M. Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton, New Jersey, 2 edition, 1971.

[170] A. H. Stone and J. W. Tukey. Generalized "sandwich" theorems. *Duke Math. J.*, 9:356–359, 1942.

[171] Betsy Stovall. Waves, spheres, and tubes: A selection of Fourier restriction problems, methods, and applications. *Notices Amer. Math. Soc.*, 66(7), Aug 2019.

[172] Robert S. Strichartz. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 1977.

[173] Robert S. Strichartz. Fourier asymptotics of fractal measures. *J. Funct. Anal.*, 89(1):154–187, 1990.

[174] V. N. Sudakov. Gaussian random processes, and measures of solid angles in Hilbert space. *Dokl. Akad. Nauk SSSR*, 197:43–45, 1971.

[175] Yukihide Tadano and Kouichi Taira. Uniform bounds of discrete Birman-Schwinger operators. *Trans. Amer. Math. Soc.*, 372(7):5243–5262, 2019.

[176] T. Tao. A sharp bilinear restrictions estimate for paraboloids. *Geom. Funct. Anal.*, 13(6):1359–1384, 2003.

[177] T. Tao and A. Vargas. A bilinear approach to cone multipliers. I. Restriction estimates. *Geom. Funct. Anal.*, 10(1):185–215, 2000.

[178] T. Tao and A. Vargas. A bilinear approach to cone multipliers. II. Applications. *Geom. Funct. Anal.*, 10(1):216–258, 2000.

[179] Terence Tao. Lecture notes: Restriction theorems and applications. Available at `https://www.math.ucla.edu/~tao/254b.1.99s/`, 1991.

[180] Terence Tao. The weak-type endpoint Bochner-Riesz conjecture and related topics. *Indiana Univ. Math. J.*, 47(3):1097–1124, 1998.

[181] Terence Tao. The Bochner-Riesz conjecture implies the restriction conjecture. *Duke Math. J.*, 96(2):363–375, 1999.

[182] Terence Tao. Spherically averaged endpoint Strichartz estimates for the two-dimensional Schrödinger equation. *Comm. Partial Differential Equations*, 25(7-8):1471–1485, 2000.

[183] Terence Tao. From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE. *Notices Amer. Math. Soc.*, 48(3):294–303, 2001.

[184] Terence Tao. On the maximal Bochner-Riesz conjecture in the plane for $p < 2$. *Trans. Amer. Math. Soc.*, 354(5):1947–1959, 2002.

[185] Terence Tao. Some recent progress on the restriction conjecture. In *Fourier Analysis and Convexity*, Appl. Numer. Harmon. Anal., pages 217–243. Birkhäuser Boston, Boston, MA, 2004.

[186] Terence Tao. Lecture notes: Fourier analysis. Available at `https://www.math.ucla.edu/~tao/247a.1.06f/`, 2006.

[187] Terence Tao. Lecture notes: Classical Fourier analysis. Available at `https://terrytao.wordpress.com/category/teaching/247b-classical-fourier-analysis/`, 2020.

[188] Terence Tao, Ana Vargas, and Luis Vega. A bilinear approach to the restriction and Kakeya conjectures. *J. Amer. Math. Soc.*, 11(4):967–1000, 1998.

[189] Peter A. Tomas. A restriction theorem for the Fourier transform. *Bull. Amer. Math. Soc.*, 81:477–478, 1975.

[190] Peter A. Tomas. Restriction theorems for the Fourier transform. In *Harmonic Analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1*, Proc. Sympos. Pure Math., XXXV, Part, pages 111–114. Amer. Math. Soc., Providence, R.I., 1979.

[191] Hans Triebel. *The Structure of Functions*, volume 97 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2001.

[192] Luis Vega. Schrödinger equations: pointwise convergence to the initial data. *Proc. Amer. Math. Soc.*, 102(4):874–878, 1988.

[193] Luis Vega. Restriction theorems and the Schrödinger multiplier on the torus. In *Partial Differential Equations with Minimal Smoothness and Applications (Chicago, IL, 1990)*, volume 42 of *IMA Vol. Math. Appl.*, pages 199–211. Springer, New York, 1992.

[194] Roman Vershynin. *High-Dimensional Probability*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2018. An Introduction with Applications in Data Science, With a foreword by Sara van de Geer.

[195] Xing Wang and Chunjie Zhang. Pointwise convergence of solutions to the Schrödinger equation on manifolds. *Canad. J. Math.*, 71(4):983–995, 2019.

[196] T. Wolff. Local smoothing type estimates on L^p for large p. *Geom. Funct. Anal.*, 10(5):1237–1288, 2000.

[197] Thomas Wolff. An improved bound for Kakeya type maximal functions. *Rev. Mat. Iberoamericana*, 11(3):651–674, 1995.

[198] Thomas Wolff. Recent work connected with the Kakeya problem. In *Prospects in Mathematics (Princeton, NJ, 1996)*, pages 129–162. Amer. Math. Soc., Providence, RI, 1999.

[199] Thomas Wolff. A sharp bilinear cone restriction estimate. *Ann. of Math. (2)*, 153(3):661–698, 2001.

[200] Thomas H. Wolff. *Lectures on Harmonic Analysis*, volume 29 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2003. With a foreword by Charles Fefferman and a preface by Izabella Łaba, Edited by Izabella Łaba and Carol Shubin.

[201] D. R. Yafaev. *Mathematical Scattering Theory*. American Mathematical Society, Providence, RI, 1992. General theory, Translated from the Russian by J. R. Schulenberger.

[202] D. R. Yafaev. *Mathematical Scattering Theory*, volume 158 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2010. Analytic Theory.

[203] Kenji Yajima. Existence of solutions for Schrödinger evolution equations. *Comm. Math. Phys.*, 110(3):415–426, 1987.

[204] Ruixiang Zhang. Restriction and incidence geometry. Available at `https://sites.google.com/view/ruixiang-zhang/home/teaching/math828_spring2020?authuser=0`, 2020.

[205] A. Zygmund. On Fourier coefficients and transforms of functions of two variables. *Studia Math.*, 50:189–201, 1974.

[206] A. Zygmund. *Trigonometric Series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.