COUNTING TO L^p

MICHAEL CHRIST

Lecture notes for the Instructional Conference on Combinatorial Aspects of Mathematical Analysis, ICMS, Edinburgh, April 1-2, 2002.

One preoccupation of harmonic analysis, and analysts, during recent decades has been the detailed study of estimates, in L^p spaces, for various linear operators whose properties are heavily influenced by the presence of some form of curvature, and/or which are associated to kernels singular along varieties more complicated than the diagonal. A second theme, which has attracted increasing attention in the last few years, is the L^p analysis of certain multilinear extensions of standard linear operators.

In these lectures I will discuss three particular problems, illustrating these overarching themes. A unifying feature is that elementary combinatorial techniques, which is to say, arguments that amount merely to counting points, have thus far yielded insights apparently inaccessible via Fourier transform-based techniques. It is well known that problems governed by cancellation are often subtle. The problems discussed here concern operators which are defined by integration against nonnegative kernels, yet nonethless enjoy subtleties of their own.

Through a discussion of these specific examples I hope to illustrate larger themes and interconnections.

1. Radon-like transforms

1.1. **Generalities.** Let X, Y be open subsets of Euclidean spaces. Suppose that to each $x \in X$ is associated a submanifold $\mathcal{M}_x \subset Y$, of some dimension k, along with a measure σ_x on $\mathcal{M}_x$ which is nonnegative and has a smooth density. Suppose moreover that these submanifolds and measures vary smoothly with x.

To these data we may associate an operator T, mapping compactly supported continuous functions on Y to continuous functions on x, defined by

$$Tf(x) = \int_{\mathcal{M}_x} f(y)\, d\sigma_x(y). \tag{1}$$

Examples of Radon-like transforms.

Example 1. *Let $Y = \mathbb{R}^d$, and let X be the manifold of all hyperplanes in $\mathbb{R}^d$. Take induced Lebesgue measure on the hyperplanes, and a natural invariant measure on X.*

Example 2. *Same as previous example, but consider instead the Grassmann manifold of all lines in $\mathbb{R}^d$.*

Example 3. *Take $X = Y = \mathbb{R}^d$ for some $d \geq 2$, and $\mathcal{M}_x = \{x + (t, |t|^2) : t \in \mathbb{R}^{d-1}\}$, where $(t, |t|^2)$ denotes a point in $\mathbb{R}^{d-1} \times \mathbb{R}^1$.*

Example 4. *Let $Y = \mathbb{R}^d$ for some $d \geq 2$, and $X = \mathbb{R}^d \times \mathbb{R}^+$. Define $\mathcal{M}_{z,r} = \{y \in \mathbb{R}^d : |z - y| = r\}$.*

Supported in part by NSF grant DMS-9970660.

1.1.1. *A nondegeneracy hypothesis.* One way to describe this situation more formally is to assume given a smooth submanifold $Z \subset X \times Y$ such that the two natural projections $\pi_X : Z \mapsto X$ and $\pi_Y : Z \mapsto Y$ are submersions at each point of Z. Then $\mathcal{M}_x = \{y \in Y : (x,y) \in Z\} = \pi_Y \pi_X^{-1}(\{x\})$.

This setup is symmetric. Thus one also has a dual family of submanifolds $\mathcal{M}'_y \subset X$, and an adjoint operator T^* mapping functions on X to functions on Y, which involves integrating over the dual family of submanifolds.

The manifolds $\mathcal{M}_x$ may be parametrized locally as $\mathcal{M}_x = \{\gamma(x,t) : t \in U\}$ where $U \subset \mathbb{R}^k$ is a small neighborhood of $0 \in \mathbb{R}^k$, $\gamma : X \times U \mapsto Y$ is C^∞, and the matrix $\partial\gamma/\partial t$ has full rank k at each point (x,t). Thus

$$Tf(x) = \int f(\gamma(x,t))\, \eta(x,t)\, dt$$

for some nonnegative auxiliary function $\eta \in C_0^\infty$. The dual manifolds $\mathcal{M}'_y$ may be parametrized in the same way, by some function γ^*.

Provided we work only in a sufficiently small neighborhood of a given point in $X \times Y$, T is bounded from $L^\infty(Y)$ to $L^\infty(X)$. By symmetry, T^* has the same property. By duality, T is therefore bounded from $L^p(Y)$ to $L^p(X)$ for $p = 1$. Then by interpolation, it is bounded for all $p \in [1, \infty]$.

In certain cases it is natural to consider global estimates, in which there is no localization to a small region. In some instances the set possesses some scale invariance, in which case any $L^p \mapsto L^q$ inequalities which themselves possess appropriate scale invariance will automatically hold in a global formulation, if they hold in a local version. For the sake of simplicity, I will concentrate on local formulations in these notes.

1.1.2. *Incidence matrices.* Many combinatorial problems can be formulated as follows. Let two finite universes A, B be given, together with a subset $Z \subset A \times B$. We say that points $a \in A$, $b \in B$ are *incident* if $(a,b) \in A \times B$. For example, A could be some set of points in $\mathbb{R}^d$, B some set of lines, and $(a,b) \in Z$ if and only if $a \in b$. In this setting, the Szemeredi-Trotter theorem gives an upper bound on the number of possible incidences, in terms of the cardinalities of A, B.

A problem of Erdos concerns two finite sets A, B of the same cardinality N. Suppose a set $E \subset A \times B$ has the property that there exists no set of four points $a_1 \neq a_2 \in A$ and $b_1 \neq b_2 \in B$ such that a_i is incident to b_j for all $i, j \in \{1, 2\}$. How large can $|E|$ be?

These lectures are concerned with continuum incidence problems.

1.1.3. *Smoothing estimates.* We are interested in quantifying the smoothing effect of Radon-like transforms. Among the ways in which a smoothing effect may be manifested are
(i) Higher integrability: For each $p \in (1, \infty)$ there exists $q > p$ such that T maps $L^p(Y)$ boundedly to $L^q(X)$. Having such an estimate for a single p implies it for all $p \in (1, \infty)$, by interpolation with the trivial $L^1 \mapsto L^1$ and $L^\infty \mapsto L^\infty$ estimates.
(ii) Sobolev space boosting: For each $p \in (1, \infty)$ there exists $\alpha > 0$ such that T maps $L^p(Y)$ boundedly to $L^p_\alpha(X)$, where the latter denotes the space off all functions possessing α derivatives in L^p, in the usual sense of distributions. Again, one p implies all $p \in (1, \infty)$. Existence of such an α is generally simplest to analyze for $p = 2$.

By the Sobolev embedding theorem, smoothing in the Sobolev boosting sense (ii) automatically implies smoothing in the higher integrability sense (i). Nonetheless, it is not automatic that any valid $L^p \mapsto L^q$ estimate is implied by some valid $L^p \mapsto L^p_\alpha$ estimate together with Sobolev embedding; in fact this is false for the main example discussed in these notes, even for $p = 2$.

1.1.4. *Curvature.* Not every Radon-like transform exhibits any smoothing effect. For example, regard $\mathbb{R}^d$ as $\mathbb{R}^{d-1} \times \mathbb{R}$ with coordinates $x = (x', x_d)$ and let T be associated to the mapping $\gamma((x', x_d), t) = (x' + t, x_d)$ from $\mathbb{R}^d \times \mathbb{R}^{d-1}$ to $\mathbb{R}^d$. Such an operator is smoothing of infinite order with respect to the first $d - 1$ coordinates, but has no smoothing effect at all with respect to the last coordinate. Some curvature is needed to produce smoothing.

In the equi-dimensional case where X, Y have the same dimension d, several equivalent characterizations of smoothing are known. For simplicity, I will state these under the simplifying assumption that the family of submanifolds may be parametrized by a real analytic function γ; see [15] for the C^∞ case. Assume to preclude trivial degeneracy that we are working near a point (x_0, t_0) such that $\eta(x_0, t_0) \neq 0$.

Consider the following conditions:

(i) There exists no (small piece of) real analytic manifold N which is invariant under γ, in the sense that $\gamma(N \times \mathbb{R}^k) \subset N$. (By this I mean more precisely that some open subset of $N \times \mathbb{R}^k$ is mapped into N by γ.)

(ii) Define $\gamma^*\gamma(x, s, t) = \gamma^*(\gamma(x, s), t)$, and the iterated mappings $\Gamma^j : X \times (\mathbb{R}^k)^{2j} \mapsto X$ by $\Gamma^j(x, s_1, t_1, \ldots s_j, t_j) = \gamma^*\gamma\big(\Gamma^{j-1}(x, s_1, t_1, \ldots, s_{j-1}, t_{j-1}), s_j, t_j\big)$. Then for each x, the determinant of some $d \times d$ minor of the Jacobian matrix $\partial\Gamma^d(x, s, t)/\partial(s, t)$ does not vanish to infinite order, as a function of $(s, t) \in \mathbb{R}^{kd+kd}$, at any $(s^0, t^0) \in \mathbb{R}^{kd+kd}$.

All four of the examples described above satisfy these two nonintegrability conditions. It is worth emphasizing that these conditions express a property of the *family* of manifolds $\mathcal{M}_x$, not of the individual manifolds themselves. The families of all hyperplanes and of all lines in $\mathbb{R}^d$ do satisfy these conditions.

These conditions are invariant under the natural action of the product of the group of diffeomorphisms of X with the group of diffeomorphisms of Y.

1.1.5. *Curvature is equivalent to smoothing.*

Theorem 1.1. [15] *In the real analytic case, the curvature conditions (i) and (ii), smoothing in the sense of higher integrability, and smoothing in the sense of Sobolev boosting, are all mutually equivalent.*

Other equivalent characterizations are also discussed in [15]. For a characterization in terms of the double fibration picture $Z \subset X \times Y$, and for a more refined classification of the possible degrees of curvature, see [48].

1.1.6. *Open questions.*

Question 1. *Suppose that T is smoothing to some degree. For which (p, q) does T map L^p to L^q?*

The goal here is a general theory, which takes as input some quantified form of the curvature and outputs numbers.

Question 2. *For which (p, q, α) does T map L^p to L^q_α? Here L^q_α denotes the Sobolev space of all functions possessing α derivatives in L^q.*

Question 3. *Suppose that T is smoothing to some degree. What is the structure of extremals for the $L^p \mapsto L^q$ estimates? That is, for what kind of functions is $\|Tf\|_q/\|f\|_p$ essentially as large as possible?*

Roughly speaking, I expect that the operator norm of T is maximal on functions which are (a slight generalization of) characteristic functions of subalgebraic sets (sets defined by a bounded finite number of inequalities involving polynomials of bounded degree) of bounded complexity.[1]

Question 4. *(Local smoothing problem) In $\mathbb{R}^2$, consider the averaging operator*

$$Af(x,r) = (2\pi)^{-1}\int_0^{2\pi} f(x + (r\cos(\theta), r\sin(\theta)))\,d\theta,$$

for (say) $r \in I = [1,2]$. For which p, q, α does A map $L^p(\mathbb{R}^2)$ to $L^q_\alpha(\mathbb{R}^2 \times I)$?

Question 5. *This is merely a special case of Question 2. Consider the operator defined by convolution in $\mathbb{R}^3$ with arc length measure on a finite portion of the helix $\{(t, \cos(t), \sin(t))\}$ (or the curve (t, t^2, t^3)). For which p, q and which $s > 0$ does this operator map $L^p(\mathbb{R}^3)$ to $L^q_s(\mathbb{R}^3)$? For an interesting partial result (subsequently improved upon by Tao and Vargas) see* [38].

Let $Y = \mathbb{R}^d$, let $X = G^{k,d}$ be the Grassmann manifold of all k-dimensional affine planes in $\mathbb{R}^d$, and let $T^{(d,k)} : L^1(Y) \mapsto L^1(X)$ be the associated k-plane transform. Let $X_0 = G_0^{k,d}$ be the Grassmann manifold of all k-dimensional subspaces of $\mathbb{R}^d$.

It is known [7] for which exponents $T^{(d,k)}$ maps $L^p(Y)$ to $L^q(G^{k,d})$. It does so for $(p,q) = ((d+1)/(k+1), d+1)$, and trivially for $(p,q) = (1,1)$; further exponents may be deduced from these by interpolation. It fails for all other pairs of exponents.

However, a more refined question is of greater interest. Regard $G^{k,d}$ as a fiber bundle over $G_0^{k,d}$, with fibers $\mathbb{R}^{d-k}$; let $(\theta, y) \in X_0 \times \mathbb{R}^{d-k}$ be coordinates over a coordinate patch in $G_0^{k,d}$. Consider mixed-norm estimates on (such a subset of) $G^{k,d}$:

$$\|F(\theta,y)\|_{L^q(L^r)} = \Big(\int \Big(\int |F(\theta,y)|^r\,dy\Big)^{q/r}\,d\theta\Big)^{1/q}.$$

Question 6. *(mixed-norm estimates for the k-plane transform) For which combinations of exponents does the k-plane transform $T^{(d,k)}$ map $L^p(\mathbb{R}^d)$ to $L^q(L^r)(G_0^{k,d})$?*

This is a refined version of the well-known *Kakeya problem.* Necessary conditions are that $dp^{-1} - (d-k)r^{-1} = k$ and $q \le p' = p/(p-1)$. The strongest conjecture not ruled out by simple counterexamples is that in the presence of these obvious necessary conditions, the mixed-norm estimate might hold whenever $1 \le p < d/k$. The larger the exponent p, the sharper the result, since the case $(p,q,r) = (1,1,\infty)$ is a trivial consequence of Fubini's theorem. This has been proved [7] for all $k \ge d/2$, and also for all $p \le (d+1)/(k+1)$. The ideas of Bourgain and Wolff have subsequently led to partial results for the more difficult remaining cases; see [1, 3, 52, 54] and the ICMS lectures of Tao.

I wish to emphasize that the combinatorial techniques illustrated in these notes represent only a few of those developed in recent years. For other methods, see in particular the works of Bourgain, Schlag, and Wolff listed in the bibliography below.

[1] This could certainly be disputed, in light of the existence of Kakeya sets, but those are extremals for certain mixed-norm estimates, and perhaps arise only in connection with special L^∞ endpoint estimates. We are not interested in exact constants here, so we are allowing a loss of a numerical factor, depending on p, q, r, which blows up as $r \to \infty$.

1.1.7. *Two methods.* **(A)** Consider some integral operator $Tf(x) = \int K(x,y)f(y)\,dy$, where $K \ge 0$. Since $|Tf| \le T(|f|)$, one may restriction attention to nonnegative f when studying $L^p \mapsto L^q$ estimates.

Suppose now that q is a positive integer. Then the L^q norm may be reexpressed as a multilinear form:

$$\int Tf(x)^q\,dx = \int \prod_{j=1}^{q} \int K(x,y_j)f(y_j)\,dy_j\,dx = \int \mathcal{K}(y_1,\dots,y_q)\prod_{j=1}^{q} f(y_j)\,dy_j$$

for a certain nonnegative kernel $\mathcal{K}$. This procedure sometimes makes sense, even when K is not a function but merely a singular measure, as is the case for Radon-like transforms; then $\mathcal{K}$ may be a singular measure, rather than a function.

This method was exploited by Oberlin [35, 36] in various ways, and of course by very many other authors. However, as we shall see, it is not the case that integer exponents always play a key role for Radon-like transforms, so the scope of this method is limited.

For a concrete example, consider the k-plane transform $T^{(d,k)}$, which takes a function f defined on $\mathbb{R}^d$, and constructs a function defined on a Grassmann manifold $G^{k,d}$ by integrating f over each k-dimensional affine subspace of $\mathbb{R}^d$. A key estimate is that $T^{(d,k)}$ maps $L^{(d+1)/(k+1)}$ to L^{d+1}. Drury [16, 17] showed that

$$\begin{aligned}(2)\quad &\int_{G^{k,d}} (T^{(d,k)}f)^{d+1} \\ &= c_{d,k}\int_{(\mathbb{R}^d)^{k+1}} |\det(x_0,x_1,\cdots,x_k)|^{k-d}\Big(\prod_{j=k+1}^{d}\int_{x_j\in\pi(x_0,\dots,x_k)} f(x_j)\,d\lambda_\pi(x_j)\Big)\prod_{j=0}^{k} f(x_j)\,dx_j\end{aligned}$$

where $\pi(x_0,\dots,x_k)$ denotes the k-dimensional simplex in $\mathbb{R}^d$ spanned by (generic) points $x_0,\dots,x_k \in \mathbb{R}^d$ and $|\det(x_0,x_1,\cdots,x_k)|$ denotes the k-dimensional volume of $\pi(x_0,\dots,x_k)$. Because $k-d$ is negative, the kernel $|\det(x_0,x_1,\cdots,x_k)|^{k-d}$ is infinite on a nontrivial variety, not merely on the diagonal $x_0 = x_1 = \cdots = x_k$. Moreover, the integral on the right-hand side of (2) extends over a $(k+1)\cdot d$-dimensional subset of the space $(\mathbb{R}^d)^{d+1}$, rather than over the whole $(d+1)\cdot d$-dimensional space.

According to this formula, the L^{d+1} norm of $T^{(d,k)}f$ thus tends to be large when much of the mass of f is concentrated near a set which is highly localized, or more generally, has a substantial part concentrated near some affine subspace. For an analysis of the multilinear operator associated to the right-hand side of (2), see [7].

(B) Suppose first that T is translation-invariant; T takes the special form $\int_{\mathbb{R}^k} f(x-h(t))\eta(t)\,dt$, where $h:\mathbb{R}^k \mapsto \mathbb{R}^d$. Thus $\gamma(x,t) = x - h(t)$, and $Tf = f * \mu$ for a certain finite positive measure μ. On the Fourier transform side, $\widehat{Tf}(\xi) = \hat f(\xi)\hat\mu(\xi)$ where $\hat\mu(\xi) = \int_{\mathbb{R}^k} e^{-i\xi\cdot h(t)}\eta(t)\,dt$.

According to Theorem 1.1, T is smoothing if and only if it maps L^2 to L^2_α for some $\alpha > 0$. This is equivalent to having a decay estimate $\hat\mu(\xi) = O(|\xi|^{-\alpha})$ as $|\xi| \to \infty$, for $\xi \in \mathbb{R}^d$.

By writing $\xi = r\nu$ where $|\nu| = 1$, $\hat\mu(\xi)$ is expressed as

$$\int e^{ir\phi_\nu(t)}\eta(t)\,dt$$

where $r \in \mathbb{R}^+$ is a large parameter, ν varies over the unit sphere, the phase functions $\phi_\nu(t) = \nu\cdot h(t)$ are real-valued, and $\eta \in C^\infty(\mathbb{R}^k)$ has compact support. Such integrals may be studied as $r \to \infty$ by such techniques as van der Corput's lemma, the method of stationary phase, and the method of steepest descent. A necessary and sufficient condition for a uniform decay estimate $O(r^{-\alpha})$ is that for each ν, the gradient $\nabla_t(\nu\cdot h(t))$ of the phase function should not vanish to infinite order at any point t in the support of the cutoff function η.

Although this reasoning does not apply directly to the general non-translation-invariant case, somewhat related Fourier-based arguments can nonetheless be applied, with varying degrees of success and precision. Obtaining the optimal exponent α is quite a bit more difficult than merely establishing the existence of some positive decay rate.

The proof in [15] of Theorem 1.1 is also based on L^2 arguments, orthogonality considerations, and cancellation, although it makes no overt use of Fourier integrals.

These notes are concerned with certain problems for which neither of these two methods yields satisfactory results, so I will not discuss them further.

1.2. **A specific operator.** We now turn to the detailed analysis of a specific operator. In $\mathbb{R}^d$, for any $d\ge 2$, consider the function

$$h(t)=(t,t^2,t^3,\dots,t^d),$$

and the family of submanifolds of $\mathbb{R}^d$ parametrized by $\gamma(x,t)=x-h(t)$. The associated operator is

$$Tf(x)=\int_{-1}^{1} f(x_1-t,x_2-t^2,\dots,x_d-t^d)\,dt;$$

instead of integrating with respect to a smooth, compactly supported function of t we simply integrate over a bounded closed interval. This makes no difference here, so far as $L^p\mapsto L^q$ mapping properties are concerned.

This situation is described by the submanifold $Z\subset\mathbb{R}^{d+d}$ consisting of all points (x,y) satisfying $y_j-x_j=(y_1-x_1)^j$ for all $2\le j\le d$. Z has high codimension $d-1$; from the point of view of incidence matrices, this means that incidences are quite rare. This rarity of incidences is a common feature of the three concrete problems discussed in these notes.

1.2.1. *Statement of theorem.* Define $Q_d=(p_d^{-1},q_d^{-1})$ and $Q_d'=(1-q_d^{-1},1-p_d^{-1})$, where

$$p_d=(d+1)/2 \quad\text{and}\quad q_d=\frac{d(d+1)}{2(d-1)}=\frac{d}{d-1}p_d>p_d.$$

Let $\mathcal{R}_d\subset[0,1]^2$ denote the convex hull of the four points $(0,0)$, $(1,1)$, Q_d, and Q_d'.

Theorem 1.2. *T is bounded from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ whenever $(p^{-1},q^{-1})\in\mathcal{R}_d$, except possibly at the endpoints Q_d,Q_d'. At these endpoints, T is of restricted weak type (p,q). T is unbounded whenever $(p^{-1},q^{-1})\notin\mathcal{R}_d$.*

1.2.2. *Remarks.* Writing $Tf=f*\mu$, one finds via van der Corput's lemma that $\hat\mu(\xi)=O(|\xi|^{-1/d})$ as $|\xi|\to\infty$, and no better. Thus T maps L^2 to the Sobolev space $L^2_{1/d}$, which in turn embeds into L^s where $s^{-1}=2^{-1}-d^{-2}$. For $d\ge 3$, the theorem says that L^2 is actually mapped to L^t where $t^{-1}=2^{-1}-2/d(d+1)$; clearly $t>s$. Thus even when $p=2$, the optimal target exponent L^q is not directly implied by the Fourier decay bound (except when $d=2$).

Except in low dimensions, neither q_d nor $p_d'=p_d/(p_d-1)$ are integers.[2] Thus neither of the two basic methods described above directly establishes the theorem.

[2] q_3 is an integer, and this was exploited by Oberlin [35] to establish the theorem for $d=3$. For $d=4$, Oberlin [36, 37] also obtained the theorem, by introducing a related auxiliary family of operators for which a key estimate again occurred at an integral exponent. No such trick is known for higher dimensions.

The point Q'_d is the conjugate of Q_d; since T is essentially its own transpose, boundedness at Q_d is equivalent to boundedness at Q'_d.

1.2.3. *Restricted weak type estimates.* An operator T is said to be of restricted weak type (p,q) if the usual inequality $\langle Tf, g\rangle \le C\|f\|_p\|g\|_{q'}$ holds whenever f,g are *characteristic functions of sets*. Here, as throughout these notes, $q' = q/(q-1)$.

This is equivalent to knowing that there exists $C<\infty$ such that for any measurable set E having finite measure, for any $\alpha\in\mathbb{R}^+$, $\big|\{x : |T(\chi_E)(x)| > \alpha\}\big| \le C\alpha^{-q}|E|^{q/p}$. Here, and throughout, χ_E denotes the characteristic (or indicator) function of a set E.

A basic interpolation theorem guarantees that if a linear operator T maps L^{p_j} boundedly to L^{q_j} for $j=0,1$, where $p_j,q_j\in[1,\infty]$, then it maps L^{p_θ} boundedly to L^{q_θ} for all $\theta\in(0,1)$ as well, where $p_\theta^{-1} = \theta p_0^{-1} + (1-\theta)p_1^{-1}$ and $q_\theta^{-1} = \theta q_0^{-1} + (1-\theta)q_1^{-1}$.

If T is merely of weak type (p_j,q_j) for $j=0,1$, then it still maps L^{p_θ} boundedly to L^{q_θ} for all $\theta\in(0,1)$, provided that $q_j\ge p_j$. This last relation is no restriction for us, since we are only interested here in higher integrability bounds, $q>p$.

The upshot, for us, is that in order to prove the positive part of Theorem 1.2, it suffices to prove restricted weak type (p_d,q_d). The restricted weak type for the exponents represented by the point Q'_d follows by duality. The points $(0,0)$ and $(1,1)$ correspond to trivial estimates valid simply because T is defined by convolution with a finite measure. Interpolation between these four points yields the entire region $\mathcal{R}_d$, except that for the two nontrivial vertices Q_d, Q'_d, only restricted weak type estimates are obtained.

1.2.4. *The iterated T^*T method.* If T is a bounded linear operator on a Hilbert space, then there is an identity $\|T\|^2 = \|T^*\circ T\|$ of operator norms. An idea with a long history in Fourier analysis is that $T^*\circ T$ may be simpler than T itself, so that this can be used to derive *a priori* upper bounds for operator norms which can be combined with auxiliary approximation arguments to prove the boundedness of operators. More generally, $T^*\circ T$ can sometimes be bounded in terms of T itself (or of T and T^*), yielding an *a priori* inequality $\|T\|^2 \lesssim \|T\| + \|T^*\|$ which can be used in the same way.

The *iterated* T^*T method proceeds by repeating this reduction, equating $\|T\|^4 = \|T^*T\|^2$ with $\|(T^*T)^2\|$; this can be continued to obtain $\|T\|^{2^{n+1}} = \|((T^*T)^{2^n}\|$ for any $n\ge 1$.

Now consider operators of the form $Tf(x) = \int f(\gamma(x,t))\eta(x,t)\,dt$, on $L^2(\mathbb{R}^d)$, where $\gamma : \mathbb{R}^d\times\mathbb{R}^1$ is smooth, and $\eta\in C_0^\infty$ is a cutoff function. Assuming that $x\mapsto\gamma(x,t)$ is a local diffeomorphism of $\mathbb{R}^d$ for all t in the support of η, the adjoint operator takes the form $T^*f(y) = \int f(\gamma^*(y,t))\tilde\eta(y,t)\,dt$ where $\tilde\eta\in C_0^\infty$ as well, so

$$T^*Tf(x) = \int Tf(\gamma^*(x,t))\tilde\eta(x,t)\,dt = \iint f(\gamma(\gamma^*(x,t),s))\eta(\gamma^*(x,t),s)\tilde\eta(x,t)\,ds\,dt.$$

This is generally a less singular operator than T itself, since it is defined by integration over what is usually a two-dimensional variety. Iterating the procedure, one eventually obtains an operator defined by integrating over $\mathbb{R}^N$ for some $N=2^n\ge d$. Under the curvature hypothesis of Theorem 1.1, such sufficiently high-order iterates take the form $f\mapsto\int K(x,y)f(y)\,dy$ where $K\in L^1_{\text{loc}}(\mathbb{R}^{d+d})$, and moreover K possesses at least a small fraction of a derivative in this same space. The same considerations apply if t takes values in some higher-dimensional parameter space $\mathbb{R}^k$.

As it stands, this method has no automatic generalization to L^p spaces for $p\ne 2$. In the first section of this course I will explain a variant which does work for other exponents.

1.3. **Necessary conditions and quasi-extremals.** Henceforth I will always assume that the ambient dimension d is ≥ 3.

For completeness we indicate why the condition $(p^{-1}, q^{-1}) \in \mathcal{R}_d$ is necessary for T to map L^p to L^q, and indeed even to be of restricted weak type (p,q). Let $\delta > 0$ be a small parameter and let E_δ be the ball of radius δ centered at 0. Then $(\chi_{E_\delta} * \mu)(x) \geq c\delta$ for all x such that $|x - h(t)| \leq \delta/2$ for some $t \in [-\frac{1}{2}, \frac{1}{2}]$. The set of all such points x has measure $\gtrsim \delta^{d-1}$. The restricted weak type estimate would imply that

$$\delta \cdot \delta^{(d-1)/q} \leq C\delta^{d/p}$$

for all small δ, with C independent of δ. Therefore a necessary condition is

$$1 + \frac{d-1}{q} \geq \frac{d}{p} \; . \tag{3}$$

Boundedness of T from $L^{p,1}$ to $L^{q,\infty}$ implies boundedness from the dual of $L^{q,\infty}$ to the dual of $L^{p,1}$, so a bit of algebra gives the second necessary condition

$$\frac{d}{q} \geq \frac{d-1}{p} \; . \tag{4}$$

A third necessary condition is found by taking E_δ to be the characteristic function of a rectangular parallelipiped centered at the origin, with sides parallel to the coordinate axes, having side lengths $(\delta, \delta^2, \ldots, \delta^d)$. Then $|E_\delta| = \delta^{d(d+1)/2}$. If $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $|x_j| \leq \delta^j/4$ for all $j \leq d$, and if $|t| < \delta/4$, then $x - h(t) \in E_\delta$. Thus $T(\chi_{E_\delta}) \geq \delta$ on a set of measure $\geq 4^{-d}|E_\delta|$. This leads as above to the necessary condition

$$1 + \frac{d(d+1)}{2q} \geq \frac{d(d+1)}{2p} \; . \tag{5}$$

This condition is self dual.

Equality holds in (4) and in (5) at Q_d, where $p = (d+1)/2$ and $q = pd/(d-1)$. Equality holds in (3) and (5) at the dual point $\tilde{Q}_d$. Equality holds in (5) when $p = q = \infty$, and in (3) for $p = q = 1$.

A fourth necessary condition is that $q \geq p$; indeed, for any translation-invariant operator to be bounded from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$, it is necessary that $q > p$.

1.4. **Basics of the method.**

1.4.1. *The quantities* α, β. Set $p = p_d$ and $q = q_d$. We have seen that Theorem 1.2 boils down to an estimate

$$\langle T(\chi_E), \chi_F \rangle \leq C|E|^{1/p}|F|^{1/q'} \tag{6}$$

for Borel sets $E, F \subset \mathbb{R}^d$ having finite Lebesgue measures $|E|, |F|$.

Let E, F be given; we may assume without loss of generality that both $|E|, |F|$ are nonzero. Define the quantities α, β by

$$\langle T(\chi_E), \chi_F \rangle = \alpha|F| = \beta|E|. \tag{7}$$

For a general geometric interpretation of the fundamental quantities α, β see [51].

1.4.2. *Reduction to a lower bound on* $|E|$. The usual approach to proving a restricted weak type estimate would be to let E and some $\lambda \in \mathbb{R}^+$ be given, and to establish an upper bound on the measure of the set $F := \{x : T(\chi_E)(x) > \lambda\}$, in terms of λ and $|E|$. Instead, we let F be given instead of λ, and establish the following *lower* bound on $|E|$.

Proposition 1.3. *Let $E, F \subset \mathbb{R}^d$ be Borel sets with positive Lebesgue measures. Let α, β be defined as in* (7). *Then*

$$|E| \geq c\alpha^{d(d+1)/2}(\beta/\alpha)^{d-1}. \tag{8}$$

Substituting $\beta = \alpha|F|/|E|$ and doing some algebra, one finds that this last inequality is simply a restatement of (6). So it suffices to establish this lower bound on $|E|$.

The proof is a counting argument. We will systematically list points of (a subset of) E, then will count them (in the sense of Lebesgue measure), paying attention to the usual issue of multiplicity of enumeration which arises in such arguments for finite sets.

When $\alpha = 1$, we are essentially studying the following problem: We are given sets E, F with the property that for each $x \in E$, the entire curve $\{x + (t, t^2, \cdots, t^d) : |t| < 1\}$ lies in F. We seek a lower bound on $|E|$ in terms of some power of $|F|$.

Before proceeding, though, I'd like to explain in more detail the connection between the lower bound (8) and the more standard form $\alpha^{q_d}|F| \leq C|E|^{q_d/p_d}$ of the restricted weak type estimate. The latter may be rewritten as $|E| \geq C^{-1}\alpha^{(d+1)/2}|F|^{(d-1)/d}$, so one might imagine that the right-hand side of (8) is equal to $\alpha^{(d+1)/2}|F|^{(d-1)/d}$. This is not the case; substituting $\beta = \alpha|F|/|E|$ transforms the right-hand side of (8) to $c\alpha^{d(d+1)/2}|F|^{d-1}|E|^{1-d}$, which quantity is not $\geq c\alpha^{(d+1)/2}|F|^{(d-1)/d}$ for arbitrary values of $\alpha, |E|, |F|$. However, if $|E|$ is $\ll \alpha^{(d+1)/2}|F|^{(d-1)/d}$, then $\alpha^{d(d+1)/2}|F|^{d-1}|E|^{1-d}$ is indeed $\gg \alpha^{(d+1)/2}|F|^{(d-1)/d}$, which is a contradiction.

1.4.3. *Generating points of* E, F. Since $|F|^{-1}\int_F T(\chi_E) = \alpha$, there exists at least one point $x \in F$ such that $T(\chi_E)(x) \geq \alpha$. This means precisely that $\big|\{t \in [-1,1] : \gamma(x,t) \in E\}\big| \geq \alpha$. Thus there exist $x \in F$ and a subset $\Omega_1 \subset [-1,1]$ such that $|\Omega_1| \geq \alpha$ and $\gamma\big(\{x\} \times \Omega_1\big) \subset E$. In this way one constructs a subset of E, but one which has Lebesgue measure zero, and in fact is one-dimensional.

To go further, we observe that the situation is symmetric with respect to the pair (E, F). Therefore there exists at least one point $y \in E$ for which $T^*(\chi_F)(y) \geq \beta$. If we are lucky enough that each point of $\gamma\big(\{x\} \times \Omega_1\big)$ has this property, then each such point gives rise to a one-dimensional subset of F, and by taking the union of these we obtain a *two*-dimensional subset of F. By iterating this procedure d times, we may hope to construct a full d-dimensional subset of either E or F (depending on the parity of d). This subset would be parametrized by $\Gamma^d(x, t_1, t_2, \ldots, t_d)$, where $(t_1, \ldots, t_d)$ ranges over a certain subset of $[-1,1]^d$, and Γ^d is a d-fold alternating iterate of γ, γ^*. This construction is clearly related to the iterated T^*T method.

1.4.4. *Parameter space towers.* For each $k \geq 1$ define $\Phi_k : [-1,1]^k \to \mathbb{R}^d$ by

$$\Phi_k(t) = h(t_1) - h(t_2) + h(t_3) - \cdots + (-1)^{k+1}h(t_k) \ ,$$

where $t = (t_1, \ldots, t_k)$. Thus $\Gamma^k(y,t) = y - \Phi_k(t)$.

Lemma 1.4. *Let E, F, α, β be as above, and let $N \geq 1$ be given. Then there exist $c_0 \in \mathbb{R}^+$ depending only on d, N, a point $y_0 \in E$, and sets $\Omega_k \subset [-1,1]^k$ for $1 \leq k \leq N$, satisfying*

$$\Omega_{k+1} \subset \Omega_k \times \mathbb{R}$$
$$|\Omega_1| \geq c_0\beta$$

such that for each odd k and $t \in \Omega_k$,

$$|\{s \in [-1,1] : (t,s) \in \Omega_{k+1}\}| \geq c_0\alpha \ ,$$
$$y_0 + \Phi_{k+1}(\Omega_{k+1}) \subset E \ ,$$

and such that for each even k and $t \in \Omega_k$,

$$|\{s \in [-1,1] : (t,s) \in \Omega_{k+1}\}| \geq c_0\beta \ ,$$
$$y_0 + \Phi_{k+1}(\Omega_{k+1}) \subset F.$$

By a *parameter space tower* we mean a sequence of sets $(\Omega_1, \Omega_2, \dots)$ satisfying all these conditions, *except* the inclusions $y_0 + \Phi_{k+1}(\Omega_{k+1}) \subset E$ and $y_0 + \Phi_{k+1}(\Omega_{k+1}) \subset F$, for some parameters $\alpha, \beta > 0$. Thus there is no link between a parameter space tower, the sets E, F, and the mappings Φ_k. This notion continues to makes sense when the variables t_j range over spaces of arbitrary dimensions. This part of the construction is very general; it applies to generalized Radon transforms without any dimensional or curvature hypotheses.

1.4.5. *Proof of the lemma.* If it were true that $T(\chi_E)(x) \geq c\alpha$ for every point $x \in F$, and also that $T(\chi_F)(y) \geq c\beta$ for every $y \in E$, then one could choose any $y_0 \in E$. Define Ω_1 to be the set of all t_1 such that $y_0 + h(t_1) \in F$. Define Ω_2 to be the set of all (t_1, t_2) such that $t_1 \in \Omega_1$, and $y_0 + h(t_1) - h(t_2) \in E$. Define Ω_3 to be the set of all (t_1, t_2, t_3) such that $(t_1, t_2) \in \Omega_2$ and $y_0 + h(t_1) - h(t_2) + h(t_3) \in F$. Continue in this way to define all Ω_k by induction.

To make this construction work in the general case, it is necessary not only to choose y_0 so that there are many values of t_1 for which $y_0 + h(t_1) \in F$, but also so that for each of these t_1, there are many t_2 such that $y_0 + h(t_1) - h(t_2) \in E$, and so on.

We accomplish this by progressively replacing E, F by smaller subsets with more favorable properties. Define

$$F_1 = \{x \in F : T(\chi_E)(x) \geq \tfrac{1}{2}\alpha\}.$$

Since

$$\langle T(\chi_E), \chi_{F \setminus F_1}\rangle \leq \int_{F \setminus F_1} \tfrac{1}{2}\alpha \, dx \leq \int_F \tfrac{1}{2}\alpha \, dx = \tfrac{1}{2}\langle T(\chi_E), \chi_F\rangle,$$

we have

$$\langle T(\chi_E), \chi_{F_1}\rangle \geq \tfrac{1}{2}\langle T(\chi_E), \chi_F\rangle.$$

Thus in a certain sense F_1 is a large set, although we are not asserting any lower bound for its Lebesgue measure $|F_1|$ in terms of $|F|$.

Now replace the pair (E, F) by (E, F_1), and reverse the roles of the two sets. Define

$$E_1 = \{y \in E : T^*(\chi_{F_1})(x) \geq \tfrac{1}{4}\beta\}.$$

By the same reasoning as in the preceding paragraph, one finds that

$$\langle T(\chi_{E_1}), \chi_{F_1}\rangle \geq \tfrac{1}{4}\langle T(\chi_E), \chi_F\rangle.$$

This may be repeated to successively construct $F_2, E_2, F_3, E_3, \dots, F_N, E_N$, losing only a factor of 2 with each step.

To prove the lemma, begin with a point $y_0 \in E_N$ for which $T(\chi_{F_{N-1}}(y_0) \geq 2^{-2N}\alpha$. Take Ω_1 to be the set of all t_1 such that $y_0 + h(t_1) \in F_{N-1}$; the choice of y_0 guarantees that the set of all such t_1 has Lebesgue measure $\geq 2^{-2N}\alpha$. Because each point $y_0 + h(t_1)$ belongs not merely to F, but to F_{N-1}, there exists a sufficiently large set of values of t_2 for which $y_0 + h(t_1) - h(t_2)$ belongs to E_{N-1}, and so on.

1.4.6. *Reduction to Jacobian estimates.* For simplicity I will assume that the dimension d is even. Letting Ω_d be as in Lemma 1.4, we know that the set E contains $\Phi_d(\Omega_d)$. Thus we seek a lower bound for the measure of this subset.

Lemma 1.5. *For each d there exists $c_d > 0$ such that*

$$\big|\Phi_d(\Omega_d)\big| \geq c_d \int_{\Omega_d} \prod_{i<j\leq d} |t_i - t_j|\, dt \tag{9}$$

Here $t = (t_1, t_2, \dots, t_d) \in [-1,1]^d \supset \Omega_d$.

If $\Phi_d : [-1,1]^d \mapsto \mathbb{R}^d$ were injective, then we would have $\big|\Phi_d(\Omega_d)\big| = \int_{\Omega_d} |J(t)|\, dt$ where

$$J(t) = \det(\partial\Phi_d/\partial t) \text{ is the Jacobian determinant of } \Phi_d.$$

Although Φ_d is not injective, nor even finite-to-one, it is generically finite-to-one, and better:

Fact. *The set of points in $\mathbb{R}^d$ which have more than $d!$ preinverses under Φ_d has Lebesgue measure zero.*

This is a direct consequence of Bezout's theorem; see [9] for further details.[3] Given this fact, it is straightforward to show that

$$\big|\Phi_d(\Omega_d)\big| \geq \frac{1}{d!} \int_{\Omega_d} |J(t)|\, dt.$$

Finally, $J(t) = c\prod_{i<j\leq d}(t_i - t_j)$.

1.4.7. *A crude lower bound.* We continue to assume that d is even; the necessary modifications for odd dimensions are uninteresting and can be found in [9].

To analyze $\int_{\Omega_d} \prod_{i<j\leq d} |t_i - t_j|\, dt$ we integrate first with respect to t_d, then t_{d-1}, and so on. Choose a sufficiently small constant $c_1 > 0$. Then for any $t' = (t_1, \dots, t_{d-1}) \in \Omega_{d-1}$,

$$\big|\{t_d : (t', t_d) \in \Omega_d \text{ and } |t_d - t_j| \geq c_1\alpha \text{ for all } j < d\}\big| \geq c\alpha.$$

For any t_d in this favorable set,

$$\prod_{j<d} |t_j - t_d| \geq c\alpha^{d-1}.$$

Thus

$$\int_{\Omega_d} \prod_{i<j\leq d} |t_i - t_j| \geq c\alpha^d \int_{\Omega_{d-1}} \prod_{i<j\leq d-1} |t_i - t_j|.$$

[3]Bezout's theorem is not directly applicable for more general Radon-like transforms, parametrized by merely C^∞ functions rather than by polynomials. However, the C^∞ case can be reduced to the polynomial case by a procedure of localization, approximation and Taylor expansion, which sacrifices merely an arbitrarily small power of $\alpha\beta$, hence merely loses a potential endpoint q in the $L^p \mapsto L^q$ final conclusion, for each p. Such endpoint results may conceivably be false in general, in the C^∞ case. I was also convinced at one time that I could prove a substitute result for C^ω mappings, which gave the same generic bounded-to-one conclusion (for real analytic mappings depending real analytically on finitely many parameters), but the full details of that argument have not been written out.

Iterating this reduction leads eventually to the conclusion

$$\int_{\Omega_d} \prod_{i<j\le d} |t_i - t_j| \ge c\alpha^{d(d+1)/2}(\beta/\alpha)^\rho \tag{10}$$

where $\rho = \rho_d = (d-1) + (d-3) + \cdots + 3 + 1$.

Putting everything together, we deduce that

$$|E| \ge c\alpha^{d(d+1)/2}(\beta/\alpha)^{\rho_d}. \tag{11}$$

This is the estimate (8) which we seek, except that we wish β/α to be raised to the power $d-1$, rather than ρ. Now $\rho = d-1$ for $d = 2$, but $\rho > d-1$ for all even $d > 2$. For $\beta \gtrsim \alpha$, (10) of course implies (8); thus we may assume henceforth that $d > 2$ and $\beta \le \alpha$.

This argument is combinatorial in the sense that it proceeds by listing points of a set (a subset of E is "listed" via the function Φ with index set Ω), then counting those points (in terms of Lebesgue measure rather than cardinality), taking multiplicity properly into account (via both the Jacobian bound and Bezout's lemma).

1.4.8. *Exercise.* Use these techniques to prove that the x-ray transform is of restricted weak type from L^p to $L^q(L^r)$ in $\mathbb{R}^d$, with $p = (d+1)/2$ and $q = r = d+1$. Explain how this argument is simply a continuum version of the "bush" construction of Bourgain.

1.5. **Comments on failure.**

1.5.1. *ρ is in optimal, in one sense.* The crude bound (10) cannot be improved without the introduction of further information. Consider the case where $\alpha = 1$ and β is very small. Let the parameter space tower Ω have the following structure: For each odd j, t_j varies over the interval $[0, \beta]$, while for even j, t_j varies over the entire unit interval $[0, 1]$. One finds that (10) is sharp, up to a bounded numerical factor. However, this example is "unphysical", in the sense that if one tries to construct actual sets E, F to which it corresponds, one finds that they are necessarily larger than the lower bound (11) indicates; the set E is actually much larger than its subset $\Phi_d(\Omega_d)$. Indeed, we will show below that the higher-order iterated set $\Phi_{2d-2}(\Omega_{2d-2}) \subset E$ always has Lebesgue measure $\gtrsim$ the desired quantity $\alpha^{d(d+1)/2}(\beta/\alpha)^{d-1}$.

The better estimate $|\Phi_d(\Omega_d)| \sim \alpha^{d(d+1)/2}(\beta/\alpha)^{d-1}$ arises when Ω has the following rather different structure: Assume $\beta \ll \alpha$. For j even, t_j varies over the entire unit interval $[0, 1]$. t_1 varies over $[0, \beta]$, while for each odd $j \ge 3$, t_j varies over $[t_{j-1} - \beta, t_{j-1} + \beta]$. Then for each odd $j \ge 3$, integration with respect to t_j leads to a factor $\sim \beta^2$, rather than $\beta^{(j+1)/2}$ as in the preceding setup.

1.5.2. *Instability.* Imagine forming the higher-order iterates Φ_k for all $k \le 2d$, with corresponding parameter space towers Ω_k. Suppose that Ω_{2d} has the same structure as described above, consisting of all $(t_1, t_2, \ldots, t_{2d})$ such that $t_j \in [0, 1]$ when j is even, and $t_j \in [0, \beta]$ when j is odd. Then $E \supset \Phi_{2d}(\Omega_{2d})$, and in particular, E contains the subset $\Phi_{2d}(\Omega')$ where $\Omega' = \{t \in \Omega_{2d} : t_j = 0 \text{ for all odd } j\}$. Thus Ω' is again a d-dimensional parameter space. The Jacobian determinant of the restriction of Φ_{2d} to Ω' is a dimensional constant times $\prod_{i<j}(t_i - t_j)$, where the product is taken over all distinct *even* $i, j \le 2d$. From this it follows that

$$|E| \ge |\Phi_{2d}(\Omega')| \gtrsim 1$$

(since $\alpha = 1$). This is a far better bound than the (optimal) lower bound $|\Phi_d(\Omega_d) \gtrsim \beta^\rho$.

We view this as a lack of stability: for some integer $k \in [d, 2d-2]$, $|\Phi_k(\Omega_k)|$ is much smaller than $|\Phi_{k+2}(\Omega_{k+2})|$. Thus the "unphysical" example is unstable in this essential sense.

The analysis described above breaks down for small variants of this example. It could be that Ω_d has the bad structure described, but Ω_{2d}, and already Ω_{d+2}, do not. In such a case one can try to obtain an improvement by choosing a favorable point $s \in \Omega_2 \subset [-1,1]^2$ and replacing Ω_d by $\{t = (t_1, \dots, t_d) \in [-1,1]^d : (s,t) \in \Omega_{d+2} \subset [-1,1]^{2+d}\}$.

Or it could be that the points t_j, for different values of $(t_1, t_3, \dots, t_{2d-1})$, vary over mutually disjoint intervals. In such a case it is less clear how to find a suitable set Ω'.

1.6. **A more refined analysis.**

1.6.1. *Band structure.* In this subsection we describe an organizational principle which makes it possible to deal systematically with the instability phenomenon discussed above, for the particular operators under discussion here. The construction involves an integer $k \in \{d, d+1, \dots, 2d-2\}$, a set $\omega \subset \Omega_k$ of favorable parameter values, and a partition of the set of indices $\{1, 2, \dots, k\}$ into nonempty disjoint subsets, called *bands.* I emphasize that this construction depends on the sets Ω_j.

Throughout the discussion we assume that $\beta \le c\alpha$ for a fixed small constant c. The basic qualitative properties of are a band structure are the following.

- No band will contain more than one even index, and no even index will belong to the same band as the index 1.
- Given a band structure, each index $j \in \{1, 2, \dots, k\}$ will be designated as belonging to exactly one of the three categories *free*, *quasi-free*, or *bound*, according to the following rules. If a band contains an even index or 1, then that index is designated as free. For bands containing neither an even index nor 1, the smallest index in the band is designated as free. Thus each band contains exactly one free index.
- Each remaining index is either quasi-free or bound. If a band $\mathcal{B}$ contains exactly two indices, then the larger of the two is quasi-free. Otherwise all remaining indices are bound.
- Each bound index *is bound to* the unique free index belonging to its band. Each quasi-free index *is quasi-bound to* the unique free index in its band, but is not said to be bound. Free indices are neither bound nor quasi-bound. The number of indices bound to any free index is thus either zero, or at least two.

Heuristically, two indices i, j will belong to the same band if $|t_i - t_j|$ is small for a comparatively large set of values of (t_i, t_j), whereas they will belong to different bands if $|t_i - t_j|$ is large for many values. Here"large" and "small" must be interpreted relative to α or β, as appropriate. Moreover, there is necessarily a gray area at the boundary between the cases of small and large separations. To deal with this gray region we will associate to ω a pair (δ, δ') of small positive real parameters; the set of all such pairs which arise will be finite and will depend only on d.

In order to be useful in the analysis, we need our band structure to have certain quantitative properties, summarized in the following lemma.

Lemma 1.6. *Let d be fixed. There exists $\delta_0 > 0$, depending only on d, such that for any parameter space (α, β) tower there exist an integer $k \in [d, 2d-2]$, a band structure, a subset $\omega \subset \Omega_k$, and parameters $\delta, \delta' \in [\delta_0, 1]$ such that for every $t \in \omega$:*

(i) *The number of free indices plus the number of quasi-free indices equals d.*

(ii)
$$|\omega| \gtrsim \begin{cases} \alpha^{k/2}\beta^{k/2} & \textit{if } k \textit{ is even} \\ \alpha^{(k-1)/2}\beta^{(k+1)/2} & \textit{if } k \textit{ is odd.} \end{cases}$$

(iii) $|t_i - t_j| \ge \delta\alpha$, *unless* i, j *belong to the same band.*
(iv) *If* i *is bound to* j *then* $|t_i - t_j| \le \delta'\alpha$.
(v) *If* i *is quasi-bound to* j *then* $c\beta \le |t_i - t_j| \le \delta\alpha$.
(vi) δ' *is sufficiently small, relative to* δ, *to satisfy the hypothesis of Lemma 1.7 below.*

This is elementary; for a proof see [9]. Related separation properties have been systematically exploited in [51].

The statement is a bit lengthy, but its content is simple. Because each t_j varies over a set of measure α whenever j is even, there will always be a large subset of Ω_d on which $|t_i - t_j| \ge c\alpha$ for all even $i \ne j$. Likewise $|t_i - t_j| \ge c\beta$ for arbitrary $i \ne j$, on at least half of Ω_d. As the example discussed above shows, difficulties begin to arise whenever there exist two or more odd indices i, j for which $|t_i - t_j|$ is small on a large part of Ω_d; here small means $\lesssim \alpha$. A more careful analysis shows that real danger arises only when there is a cluster $\mathcal{I}$ of three or more indices for which $|t_i - t_j| \lesssim \alpha$ for all $i \ne j \in \mathcal{I}$, for all points in a significant fraction of Ω_d. When this happens, we iterate twice more to produce two new variables, t_{d+1} and t_{d+2}. Since $d + 2$ is even, t_{d+2} is automatically $c\alpha$–separated from all other variables on a significant proportion of Ω_{d+2}. The other new variable, t_{d+1}, may or may not have satisfactory separation properties. We continue iterating until we've produced exactly d indices with satisfactory separation properties. The bands are the clusters occurring in this discussion. Bands with one or two indices are good, while those with three or more contain only one index j for which the coordinates t_j are useful for our purpose.

Suppose that all of the conclusions of the lemma are satisfied. Define M to be the number of quasi-free indices. Then M satisfies

$$\begin{cases} M + \frac{k}{2} & \le d - 1 & \text{if } k \text{ is even} \\ M + \frac{k-1}{2} & \le d - 1 & \text{if } k \text{ is odd.} \end{cases} \tag{12}$$

Indeed, if k is even then the number N of free indices is at least $1 + \frac{k}{2}$, since all even indices are free and 1 is free. For odd k the same reasoning gives $N \ge \frac{k+1}{2}$. Since $M = d - N$, the claim follows.

1.6.2. *The slicing method.* In the remainder of the argument we show how the setup of Lemma 1.6 leads to a good lower bound for $|E|$. Let $\Lambda = \{m_1, \dots, m_d\}$ be the set of all indices that are free or quasi-free, and define $\tau = (t_{m_1}, \dots, t_{m_n}) \in \mathbb{R}^d$. For each bound index $i \le k$ let $B(i)$ be the unique free index to which i is bound, and set

$$s = (t_i - t_{B(i)})_{i \notin \Lambda} \in \mathbb{R}^{k-d} .$$

The map $t \mapsto (\tau, s)$ is a diffeomorphism; we regard t as a function of (τ, s) and introduce

$$G_s(\tau) = \Phi_k(t(\tau, s)).$$

For any $s \in \mathbb{R}^{k-d}$, consider the slice $\omega_s = \omega \cap \{t(\tau, s) : \tau \in \mathbb{R}^d\}$. Assume that k is even, so that E contains $\{G_s(\tau) : t(\tau, s) \in \omega_s\}$. Thus

$$|E| \ge |\{G_s(\tau) : t(\tau, s) \in \omega_s\}|. \tag{13}$$

By Bezout's theorem, as in the proof of (10), for each $s \in \mathbb{R}^{k-d}$ any generic point in $\mathbb{R}^d$ has at most $d!$ preimages under G_s. Thus

$$|E| \geq (d!)^{-1} \int_{\{\tau : t(\tau,s)\in\omega_s\}} \big|\det(\partial G_s/\partial\tau)(\tau,s)\big| d\tau$$
$$\gtrsim \min\{|\det(\partial G_s/\partial\tau)(\tau,s)|\} \cdot |\omega_s|,$$

where the minimum is taken over all τ such that $t(\tau,s) \in \omega_s$, and of course $|\omega_s|$ denotes the d-dimensional measure of the slice. Thus a lower bound on $|E|$ can be obtained by choosing a favorable parameter s and obtaining lower bounds both on the Jacobian determinant of G_s with respect to τ, and on the measure of the slice.

1.6.3. *A Jacobian bound.* For the Jacobian determinant of Φ_d we had the explicit Vandermonde formula. By treating G_s as a perturbation of a mapping $\tau \mapsto \pm\tau_1 \pm \tau_2 \pm \cdots \pm \tau_d$, one can derive a related lower bound:

Lemma 1.7. *For each $\delta \in \mathbb{R}^+$, there exist $\delta', c \in \mathbb{R}^+$ such that for all indices k and parameter sets ω possessing the structure described in Lemma 1.6, for each point (τ, s) such that $t(\tau,s) \in \omega$,*

$$\big|det(\partial G_s/\partial\tau)\big| \geq c\alpha^{d(d-1)/2}(\beta/\alpha)^M$$

where M is the number of quasi-free indices associated to ω. Here c depends only on d, δ, not on α, β.

See [9] for a proof.

1.6.4. *Conclusion of proof.* With the aid of Lemma 1.7, the proof of the main inequality (8) may be concluded as follows, under the additional assumption that k is even. As above, we conclude that

$$|E| \gtrsim |\omega_s| \cdot \alpha^{d(d-1)/2}(\beta/\alpha)^M.$$

Now E contains the union of all such slices, but we will exploit this in a very weak way by remarking that $|E|$ is at least as large as the *average* of the measures of the images of all slices. We have $|s| \leq \delta'\alpha$ by condition (iv) of Lemma 1.6, whenever $t(\tau,s) \in \omega$. Therefore integrating over the ball $B = \{s \in \mathbb{R}^{k-d} : |s| \leq \delta'\alpha\}$ on both sides of the preceding inequality yields

$$|B| \cdot |E| \gtrsim \alpha^{d(d-1)/2}(\beta/\alpha)^M|\omega| \ ,$$

whence by the lower bound (ii) on $|\omega|$,

$$|E| \gtrsim \alpha^{d(d+1)/2}(\beta/\alpha)^{M+\frac{k}{2}} \ .$$

Since $\beta \leq \alpha$ and $M + (k/2) \leq d - 1$, this gives (8). □

The case of odd k is quite similar, but will not be treated in these notes.

1.6.5. *A comment.* The band structure setup described above is a special device, adapted to the particular operator under discussion. To what extent it generalizes, or what form such a generalization should take, I do not know. The slicing argument, in contrast, is rather general. I believe that stability considerations are likewise of fundamental importance in great generality.

1.6.6. *Quasi-extremals revisited.* By a quasi-extremal for an operator T we mean a set E which is one of an ordered pair (E, F) of sets for which $\langle T(\chi_E), \chi_F\rangle/\|f\|_p\|g\|_q$ is as large as possible, *up to a bounded multiple*, for some (p, q) on the boundary of the set of all pairs of exponents for which L^p is mapped to L^q by T. This is an informal notion; we are not interested in the exact best constant in such an inequality.

We noted earlier that necessary conditions on those pairs (p, q) for which the operator maps L^p to L^q can be derived by examining certain examples, described above. Those quasi-extremals may seem rather natural, and are not difficult to guess. Nonetheless, the above proof sheds further light on the structure of those quasi-extremals in two ways. Firstly, it constructs them, essentially as images $\Phi_d(\Omega_d)$ for particular parameter space towers Ω_d; it indicates which Ω_d yield quasi-extremals. Secondly, it produces a *larger* family of quasi-extremals; these depend on two independent dimensional parameters α, β (whereas the quasi-extremals described above fall into three families, each depending on a single parameter).[4] Thirdly, it indicates that, very roughly speaking, any quasi-extremal is one of these.

These quasi-extremals come in three families. Assume for simplicity that d is even. For any $0 < \beta \le \alpha \in (0, 1]$, the basic inequality $|E| \gtrsim \alpha^{d(d+1)/2}(\beta/\alpha)^{d-1}$ is "saturated" when Ω_d consists of all $t \in [-1, 1]^d$ for which $|t_j| \le \alpha$ for j even, $|t_1| \le \beta$, and $|t_j - t_{j-1}| \le \beta$ for all odd $j \ge 3$. Then $\Phi_d(\Omega_d)$ is essentially

$$\{x = (x_1, \dots, x_d) \in \mathbb{R}^d : |x_1| \le \alpha \text{ and } |x_j + (-1)^j x_1^j| \le C\beta\alpha^{j-1} \text{ for all } 2 \le j \le d\}.$$

Corresponding to these are certain dual sets F lying in the target space $\mathbb{R}^d$. For $\beta \sim \alpha$, $\Phi_d(\Omega_d)$ is essentially a Euclidean ball of radius $\sim \alpha$ centered at the origin (or of course at an arbitrary point in $\mathbb{R}^d$).

A general description of these sets by Tao and Wright [51] characterizes them in the double fibration picture as projections of two-parameter families of "balls" lying in $\mathbb{R}^{d+1}$.

1.6.7. *A generalization.* Theorem 1.2 has been vastly extended by Tao and Wright [51] in a beautiful paper. Their proof uses some of the same ingredients, and introduces some new ideas; see Lemma 2.3 for one of them. For various reasons, I have nonetheless decided to focus in these lectures on this concrete operator, and on the older proof.

1.7. **A higher-dimensional variant.** Define $h : \mathbb{R}^{d-1} \mapsto \mathbb{R}$ to be $h(t) = |t|^2$. Consider the Radon-like transform $Tf(x) = \int_{|t|<1} f(x - h(t))\,dt$. Here $f : \mathbb{R}^d \mapsto \mathbb{C}$, while $t \in \mathbb{R}^{d-1}$.

Proposition 1.8. *T is of restricted weak type $(\frac{d+1}{d}, d+1)$.*

T also maps L^p boundedly to itself for all $p \in [1, \infty]$. From interpolation it follows that T maps L^p to L^q whenever (p^{-1}, q^{-1}) belongs to the closed triangle with vertices $(1, 1), (0, 0), (\frac{d}{d+1}, \frac{1}{d+1})$, except possibly at the third vertex. It is unbounded for all pairs (p^{-1}, q^{-1}) outside the triangle, as is readily seen by computing the actions of T (and of its transpose, which is essentially the same as T) on characteristic functions of small balls and on rectangles with sidelengths (δ, δ^2); this is left as an exercise for the reader.

However, this result is not the best possible; the operator is actually of strong type $(\frac{d+1}{d}, d+1)$, as follows from the Fourier transform-based method outlined above. The basic estimate is that T maps L^2 to the Sobolev space $L^2_{(d-1)/2}$; this is combined with an interpolation argument to an appropriate analytic family of deformations. More combinatorial

[4] Of course, any translate of a quasi-extremal is likewise a quasi-extremal, yielding an additional d parameters not included in this discussion.

proofs have been obtained independently by Schlag [45] and this author. The purpose of this discussion is merely to further illustrate the counting method.

The method discussed in §1.4 applies here, with two twists. Let E, F be arbitrary measurable subsets of $\mathbb{R}^d$, with finite measures. Set $X = \langle T(\chi_E), \chi_F\rangle$, and define α, β by the equations $X = \alpha|F| = \beta|E|$. We aim to prove the inequality $|E| \ge c\alpha^{d/(d-1)}\beta^{1/(d-1)}$, which translates to $X \le C|E|^{d/(d+1)}|F|^{d/(d+1)}$, the desired restricted weak type estimate.

Given some $x_0 \in \mathbb{R}^d$, define the second iterate mapping $\phi(t,s) = x_0 - h(t) + h(s)$. By Lemma 1.4, there exist a point $x_0 \in \mathbb{R}^d$ and a (β,α) parameter space tower $\Omega_1 \subset \mathbb{R}^{d-1}$, $\Omega = \Omega_2 \subset \mathbb{R}^{2(d-1)}$ such that $\phi(\Omega) \subset E$, $\Omega_2 \subset \Omega_1 \times \mathbb{R}$, $|\Omega_1| = \beta$ and for each $t \in \Omega_1$, $|\{s : (t,s) \in \Omega_2\}| = \alpha$.

The case $d = 2$ is contained in Theorem 1.2 concerning convolution with measures on the curve $(t, t^2, \cdots, t^d) \subset \mathbb{R}^d$, so we assume henceforth that $d \ge 3$. The first twist is that since ϕ maps a space of dimension $2d-2$ to one of dimension $d \ne 2d-2$, we cannot[5] obtain a lower bound for $|E|$ by integrating a Jacobian determinant of ϕ.

Instead, we employ the inflation device (see the discussion of trilinear operators in §3.7.2 below). Define

$$\Omega' = \{(t, s_1, \ldots, s_{d-1}) \in (\mathbb{R}^{d-1})^d : (t, s_j) \in \Omega \text{ for every } 1 \le j \le d-1\}.$$

For each $t \in \Omega_1$ let $\omega_t = \{s \in \mathbb{R}^{d-1} : (t,s) \in \Omega\}$. Consider the mapping $\Psi : (\mathbb{R}^{d-1})^d \mapsto (\mathbb{R}^d)^{d-1}$ defined by

$$\Psi(t, s_1, \ldots, s_{d-1}) = \begin{pmatrix} (t,s_1) & (t,s_2) & \cdots & (t,s_{d-1}) \end{pmatrix}.$$

Ψ is a mapping between two spaces of the same dimension, and $\Psi(\Omega') \subset E^{d-1}$, so if we define $J(t,s)$ to be the Jacobian determinant $\partial\Psi/\partial t\partial s_1 \cdots \partial s_{d-1}$, then since Ψ is algebraic, it follows[6] as in Lemma 1.5 above that

$$|E|^{d-1} \ge c_d \int_{\Omega'} |J(t, s_1, \ldots, s_{d-1})|\, dt\, ds_1 \cdots ds_{d-1}.$$

Substituting $s_j = t + u_j$, one finds that $J(t,s) = c_d \det(u)$ where $\det(u) = \det(u_1, \ldots, u_{d-1})$ denotes the determinant of the $(d-1)\times(d-1)$ matrix whose entries are the coordinates of the points u_j. This is the second twist: the Vandermonde products encountered in our analysis of convolution with the curves $(t, t^2, \cdots, t^d)$ have been replaced by such higher-dimensional determinants. This is no surprise; as will be seen again in the discussion of a restricted x-ray transform below in §2, the algebraic aspects of this type of analysis vary from one operator to another and of course reflect the details of the underlying geometric structure.

The following lemma concerning integrals of determinants suffices to conclude the proof.

Lemma 1.9. *Let $n \ge 1$. Then there exists $c_n > 0$ such that for any measurable set $A \subset \mathbb{R}^n$ having finite measure, Then $\int_{y=(y_1,\cdots,y_n)\in A^n} |\det(y_1, \ldots, y_n)|\, dy \ge c_n|A|^{n+1}$.*

Applying this to $A = \omega_t$, with $n = d-1$, for each $t \in \Omega_1$ gives $\int_{u\in\omega_t^{d-1}} |\det(u)|\, du \ge |\omega_t|^d$; integrating with respect to t then yields the lower bound $|E|^{d-1} \ge c\alpha^d\beta$, so $|E| \ge c\alpha^{d/(d-1)}\beta^{1/(d-1)}$ as desired.

[5]The slicing method, as outlined above, could be used, but it seems to lead only to rather poor results. This has to do with the interaction between the dimension and the rate of vanishing of the relevant Jacobian along the diagonal.

[6]Here simple linear algebra gives the needed generic uniform boundedness of sets of preimages; there is no need to apply Bezout's theorem.

Proof. The proof is based on an induction on the dimension n; for the purpose of making such an induction it is useful to prove a slightly more general statement: Let $F \in L^1 \cap L^\infty(\mathbb{R}^n)$ be nonnegative, and assume that F does not vanish identically (modulo null sets). Then

$$\int_{\mathbb{R}^{n \cdot n}} |\det(y_1, \cdots, y_n)| \prod_{j=1}^{n} F(y_j)\, dy_j \geq c_n \|F\|_1^{d+1} \|F\|_\infty^{-1}.$$

To prove this, let F be given and define $R \in (0, \infty)$ by the relation $\int_{|x|<R} F(x)\, dx = \frac{1}{2} \int_{\mathbb{R}^n} F(x)\, dx$. We now seek a lower bound for the integral over the restricted domain $\{|x| < R\}^{n-1} \times \{|x| > R\}$.

Consider an arbitrary point $y_n \in \mathbb{R}^n$ satisfying $|y_n| > R$. For any $x \in \mathbb{R}^n$ define x' to be the orthogonal projection of x onto the hyperplane orthogonal to y_n; we regard this hyperplane as being a copy of $\mathbb{R}^{n-1}$. Factor $|\det(y)| = |y_n| \cdot |\det(y_1', \cdots, y_{n-1}')|$, where the latter determinant is that of an $(n-1) \times (n-1)$ matrix. Fix y_n, and change coordinates by a rotation so that y_n is a scalar multiple of $(0, 0, \cdots, 0, 1)$; thus the hyperplane $\mathbb{R}^{n-1}$ orthogonal to y_n is identified with $\mathbb{R}^{n-1} \times \{0\}$. Upon integrating with respect to $dy_1 \cdot dy_{d-1}$, we thus obtain

$$F(y_n)|y_n| \int_{(\mathbb{R}^{n-1})^{n-1}} |\det(y_1', \cdots, y_{n-1}')| \prod_{j=1}^{n-1} \tilde{F}(y_j)\, dy_j \tag{14}$$

where $\tilde{F}(y') = \int_{\mathbb{R}} F(y', t) \cdot \chi_R(y', t)\, dt$ and χ_R denotes the characteristic function of the ball centered at the origin with radius R in $\mathbb{R}^n$. Now $\|\tilde{F}\|_{L^1(\mathbb{R}^{n-1})} = \frac{1}{2}\|F\|_{L^1(\mathbb{R}^n)}$, while $\|\tilde{F}\|_\infty \leq 2R\|F\|_\infty$. Thus by induction, (14) is greater than or equal to $cF(y_n)|y_n| \cdot \|\tilde{F}\|_1^n \|\tilde{F}\|_\infty^{-1} \geq cF(y_n)|y_n| \cdot \|F\|_1^n R^{-1} \|F\|_\infty^{-1}$. Since $|y_n| \geq R$, this is $\geq cF(y_n)\|F\|_1^n \|F\|_\infty^{-1}$. Integrating over all $|y_n| > R$ then yields the desired bound. □

An alternative argument (also observed by Schlag [45]) is based on a rearrangement inequality for the multilinear expression (2) mentioned above in our discussion of the x-ray transform; that inequality shows that if f is taken to be the characteristic function of a set A of any prescribed Lebesgue measure, the integral is minimized when A is the characteristic function of a ball; for balls the inequality is elementary. The downside of this proof is that it is unlikely to extend to many other operators (e.g. convolution with measures on surfaces whose curvature degenerates in various ways).

Problem 1.1. *Suppose that $A \subset \mathbb{R}^n$ and $|A| = 1$; by scaling, proving the lemma in this case is equivalent to proving it for sets of arbitrary measure. The restriction to A of Lebesgue measure is thus a probability measure, and likewise for A^n. The lemma thus states a lower bound on the expectation of $|\det(y)|$ over A^n, where $y = (y_1, \cdots, y_n)$. Establish the following strong bound: There exists $c_n > 0$ such that for any such A, $|\{y \in A^n : |\det(y)| \geq c_n\}| \geq c_n$.*

Problem 1.2. *Extend the nonconcentration lemma 2.3 to higher dimensions in the following form: For any $\varepsilon > 0$ and any dimension d there exists $c > 0$ such that for any set $A \subset \mathbb{R}^d$ satisfying $0 < |A| < \infty$, there exists a rectangular parallelepiped $I \subset \mathbb{R}^n$ such that $|A \cap I| \geq \frac{1}{2}|A|$, $|I| = 2^m|A|$ for some $m \in \{0, 1, 2, \ldots, \}$, and for any rectangular parallelepiped $I' \subset \mathbb{R}^n$ satisfying $|I'| = \frac{1}{2}|I|$, $|A \cap (I \backslash I')| \geq c2^{-\varepsilon m}|A|$. (Note that we have not required I' to be a subset of I.)*

Remark. *Analogous statements can be formulated with rectangular parallelepipeds replaced by other families of sets; for instance, by the collection of all subalgebraic sets of fixed degree and complexity.*

Problem 1.3. *Use the result of the preceding problem to construct a third proof of Lemma 1.9.*

2. A restricted X-ray transform

2.1. **A line complex and associated operator.** For each θ in a fixed compact subinterval I_0 of $\mathbb{R}^1$, define $\nu_\theta = (1, \theta, \theta^2, \dots, \theta^{d-1})$ and $V_\theta = \{y \in \mathbb{R}^d : y \perp \nu_\theta\}$. Equip the hyperplane V_θ with Lebesgue measure. Define the d-dimensional manifold $G = \{(\theta, y) \in I_0 \times \mathbb{R}^d : y \in V_\theta\}$.

Let f be defined on $\mathbb{R}^d$. For each $(\theta, y) \in G$, define

$$Xf(\theta, y) = \int_{|s| \le C_0} f(y + s\nu_\theta)\, ds, \tag{15}$$

where C_0 is any constant. On G we define mixed norms by

$$\|g\|_{L^q(L^r)(G)} = \Big(\int_{I_0} \big(\int_{V_\theta} |g(\theta, y)|^r\, dy \big)^{q/r} d\theta \Big)^{1/q}. \tag{16}$$

Here dy denotes Lebesgue measure on the hyperplane V_θ for each θ. We write simply $L^q(G)$ when $r = q$.

2.1.1. *Mixed norm estimates.* Our main result will characterize those exponents for which X maps $L^p(\mathbb{R}^d)$ (more precisely, L^p of any arbitrary fixed *bounded* subset of $\mathbb{R}^d$) boundedly to $L^q(L^r)(G)$, except for endpoints. Necessary conditions [18] are:

$$dp^{-1} \le (d-1)r^{-1} + 1 \tag{17}$$

$$d(d-1)p^{-1} \le 2q^{-1} + d(d-1)r^{-1}, \tag{18}$$

$$(d-2)(d+1)p^{-1} \le d(d-1)r^{-1}. \tag{19}$$

The key exponents for this problem turn out to be

$$p_0 = q_0 = \frac{d+2}{d}, \quad \text{and} \quad r_0 = \frac{d^2+d-2}{d^2-d-2} = \frac{d^2-d}{d^2-d-2} \cdot q_0. \tag{20}$$

Indeed, the set of all triples (p^{-1}, q^{-1}, r^{-1}) satisfying the three necessary conditions listed is the convex hull of the three points $(1,0,1)$, $(0,0,0)$ and $(p_0^{-1}, q_0^{-1}, r_0^{-1})$, together with all points (p^{-1}, q^{-1}, r^{-1}) for which there exists a triple $(\tilde{p}^{-1}, \tilde{q}^{-1}, \tilde{r}^{-1})$ in that convex hull satisfying $p \ge \tilde{p}$, $q \le \tilde{q}$, and $r \le \tilde{r}$.

2.1.2. *The Kakeya problem.* For information on the Kakeya problem, see the ICMS lectures of Tao.

2.1.3. *Statement of theorem.*

Theorem 2.1. *Let $d \ge 3$ be arbitrary, and let I_0, C_0 be fixed. Then X maps $L^p(\mathbb{R}^d)$ to $L^q(L^r)(G)$ whenever (p^{-1}, q^{-1}, r^{-1}) satisfies* (17), (18), *and* (19) *with* strict *inequality in each.*

We will prove a sharper restricted weak type inequality:

$$\langle g, Xf \rangle \le C\|f\|_{L^p} \|g\|_{L^{q'}(L^{r'})} \tag{21}$$

whenever f, g are characteristic functions of measurable sets, for all (p, q, r) satisfying (17), (18), and (19), where q', r' are the exponents conjugate respectively to q, r. By a crude interpolation argument, this implies boundedness from L^p to $L^q(L^r)$ except at endpoints, as stated above; the details are left to the reader.

The result is true, and is already well known, for $d = 2$, where one is dealing with the full complex of all lines in $\mathbb{R}^2$, and the critical exponent r_0 equals ∞. That case can be handled by a modification of the method outlined here, as will be indicated after the conclusion of the proof for $d \ge 3$.

2.2. **The analysis.**

2.2.1. *Reparametrizations.* It will be convenient to reparametrize G. Instead of taking V_θ to be the orthocomplement of ν_θ, we instead replace it by a fixed copy of $\mathbb{R}^{d-1}$, namely $V = \{y \in \mathbb{R}^d : y_1 = 0\}$. Thus G will be identified with a compact subset of $\mathbb{R}\times\mathbb{R}^{d-1} = \mathbb{R}\times V$. Define for $(\theta, y) \in G$ and $s \in \mathbb{R}$

$$\gamma(\theta, y, s) = (s, y_2 + s\theta, \dots, y_d + s\theta^{d-1}). \tag{22}$$

Then X may be identified with the operator

$$Xf(\theta, y) = \int_{\mathbb{R}} f(\gamma(\theta, y, s))\, ds \tag{23}$$

where it is always understood that the integration with respect to s is taken over a fixed compact subset of $\mathbb{R}$. Mixed norms are now

$$\|g\|_{L^q(L^r)(G)} = \Big(\int \big(\int_V |g(\theta, y)|^r\, dy\big)^{q/r}\, d\theta\Big)^{1/q}. \tag{24}$$

The adjoint operator defined by $\int_G Xf \cdot g = \int_{\mathbb{R}^d} f \cdot X^*g$ is

$$X^*g(x) = \int_{\mathbb{R}} g(\gamma^*(x, t))\, dt \tag{25}$$

where

$$\gamma^*(x, t) = (t, x_2 - x_1 t, x_3 - x_1 t^2, \dots, x_d - x_1 t^{d-1}) \ . \tag{26}$$

We denote the characteristic function of a set E by χ_E. For any two measurable sets $E \subset \mathbb{R}^d$, $F \subset G$, define

$$\mathcal{X}(E, F) = \langle \chi_F, X(\chi_E)\rangle = \int_F X(\chi_E). \tag{27}$$

This bilinear form may be written in a formally symmetric way as $\langle X^*(\chi_F), \chi_E\rangle$.

Fix any two such measurable sets having finite measures. Define the fundamental quantities α, β by

$$\begin{aligned} \alpha &= \mathcal{X}(E, F)/|F| \\ \beta &= \mathcal{X}(E, F)/|E| \ . \end{aligned} \tag{28}$$

The method developed in the first section of these notes, together with the Jacobian formula (31), can be used to show that

$$\begin{aligned} &|E| \ge c\alpha^d \beta^{d(d-1)/2} \text{ for } d \text{ even}, \\ &|F| \ge c\alpha^{d-1} \beta^{(d^2-d+2)/2} \text{ for } d \text{ odd}. \end{aligned} \tag{29}$$

for some fixed constant $c > 0$.

2.2.2. *Exercise.* Use (29) to prove that X maps $L^p(\mathbb{R}^d)$ to $L^q(G)$ for all (p^{-1},q^{-1}) in the convex hull of the three points $(1,1),(0,0),(p_1^{-1},q_1^{-1})$ where $p_1=d(d+1)/(d^2-d+2)$ and $q_1=(d+1)/(d-1)$ (except that at the endpoint (p_1^{-1},q_1^{-1}) itself, the method yields merely a restricted weak type estimate).

This result is optimal of its type, except for the slight gap at the endpoints. Our interest here is in more subtle mixed norm estimates, so we leave the details to the reader.

2.2.3. *Jacobian formula.* Define a sequence of maps Φ_k mapping $\mathbb{R}^k$ to $\mathbb{R}^d$ if k is even, and to G if k is odd, by fixing a point $x^0\in\mathbb{R}^d$, defining $\Phi_1(t_1)=\gamma^*(x^0,t_1)$, $\Phi_2(t_1,s_1)=\gamma(\Phi_1(t_1),s_1)$, and in general

$$\begin{aligned}\Phi_{2k+1}(t_1,s_1,t_2,s_2,\dots,t_{k+1})&=\gamma^*(\Phi_{2k}(t_1,s_1,t_2,s_2,\dots,t_k,s_k),t_{k+1}),\\ \Phi_{2k+2}(t_1,s_1,t_2,s_2,\dots,t_{k+1},s_{k+1})&=\gamma(\Phi_{2k+1}(t_1,s_1,t_2,s_2,\dots,t_{k+1}),s_{k+1}).\end{aligned}\tag{30}$$

It is a pleasant exercise to compute the Jacobian determinant $J(s,t)=\det(\partial\Phi_d/\partial(s,t))$:

$$J=c_d\cdot\begin{cases}\prod_{j=1}^D(s_j-s_{j-1})\prod_{1\le j<k\le D}(t_j-t_k)^4 & \text{for } d=2D\\ \prod_{j=1}^D(s_j-s_{j-1})\prod_{1\le j<k\le D}(t_j-t_k)^4\prod_{j=1}^D(t_j-t_{D+1})^2 & \text{for } d=2D+1\end{cases}\tag{31}$$

for some nonzero constants c_d. Here s_0 is defined to equal x_1, the first coordinate of the initial point x^0, to make this formula appear more harmonious. For the derivation of (31), and for other details omitted in these notes, see [11]. A special feature of this Jacobian, which simplifies the subsequent analysis, is that the variables s_j do not mix with the variable t_k.

2.2.4. *Converting information to mixed norms.* The main new element in this theorem is the appearance of mixed norms; lower bounds on the measures of E,F must be reexpressed in those norms. The following simple lemma will permit this.

For any set $F\subset G=\mathbb{R}^1\times\mathbb{R}^{d-1}$, denote by $\pi(F)$ its projection onto the first factor $\mathbb{R}^1$.

Lemma 2.2.

$$|F|\le\|\chi_F\|_{L^{q_0'}(L^{r_0'})}^{(d+2)(d-1)/2d}|\pi(F)|^{1/d}.\tag{32}$$

Proof. Defining $f(\theta)=|\{y\in\mathbb{R}^{d-1}:(\theta,y)\in F\}|$, we have

$$\|\chi_F\|_{L^{q_0'}(L^{r_0'})}=\Big(\int_{\pi(F)}f(\theta)^{q_0'/r_0'}\,d\theta\Big)^{1/q_0'},\tag{33}$$

and $|F|=\int_{\pi(F)}f$. An application of Hölder's inequality thus yields the lemma. □

2.2.5. *Noncentration.* The next lemma, an adaptation of an idea of Tao and Wright [51], asserts that any subset of $\mathbb{R}^1$ may be regarded as being roughly even distributed over some interval, albeit in a *very* weak sense.

Lemma 2.3. *For any $\varepsilon>0$ there exists $c>0$ such that for any Borel set $A\subset\mathbb{R}^1$ satisfying $0<|A|<\infty$, there exist $m\in\{0,1,2,\dots\}$, and a closed bounded interval $I\subset\mathbb{R}$ such that $|I|=2^m|A|$ and for any subinterval $I'\subset I$ of length $\frac12|I|$,*

$$|A\cap(I\backslash I')|\ge c2^{-\varepsilon m}|A|.\tag{34}$$

This says that $A\cap I$ is a reasonably large part of A, which is not too highly concentrated in I. However, this notion of noncentration is an extremely weak one, which permits $A\cap I$ to lie in the union of *two* subintervals I_1,I_2 of I, for which $|I_j|/|I|$ is arbitrarily small, provided that the distance from I_1 to I_2 is comparable to the length of I. See [51] for further information on this notion. Related ideas had earlier been exploited by Bourgain and by Wolff.

The proof [11] is a very simple stopping-time argument. Begin with a large interval I_M such that $|A \cap I_M| \ge \frac{1}{2}|A|$, with $|I_M| = 2^M|A|$ for some $M \in \mathbb{N}$. If I_M does not have the required property, then there is some subinterval $I_{M-1} \subset I_M$ of length $\frac{1}{2}|I_M|$ such that $|A \cap I_{M-1}| \ge (1 - 2^{-M\varepsilon})|A \cap I_M|$. Repeat this process for I_{M-1}, with $2^{-M\varepsilon}$ replaced by $2^{-(M-1)\varepsilon}$. Continuing in this way, and using the convergence of the infinite product $\prod_{k=0}^{\infty}(1 - 2^{-k\varepsilon})$, one reaches the desired conclusion after at most M subdivisions, with a constant $c > 0$ independent of M. □

In the application, the loss of $2^{-m\varepsilon}$ will be counterbalanced by a gain of a factor of 2^m when $|I| \sim 2^m|A|$. Variants of this lemma are of course possible.

2.2.6. *Localization.* To prove Theorem 2.1, it suffices to establish the restricted weak type bound at the endpoint $(p_0^{-1}, q_0^{-1}, r_0^{-1})$: $\mathcal{X}(E,F) \le C|E|^{1/p_0}\|\chi_F\|_{L^{q_0'}(L^{r_0'})}$ for any measurable sets $E \subset \mathbb{R}^d$, $F \subset G$ having finite measure. Moreover, it suffices to prove this for any F and any $\beta > 0$, with

$$E = \{x \in \mathbb{R}^d : \beta \le X^*(\chi_F)(x) < 2\beta\}. \tag{35}$$

For each $x \in E$ choose a set $A_x \subset \mathbb{R}$ of measure β such that $\gamma^*(x,t) \in F$ for all $t \in A_x$. Then choose an interval $I_x \subset \mathbb{R}$ which satisfies the conclusions of Lemma 2.3 relative to the set A_x, so that $|I_x| = 2^m\beta$ for some integer $m \ge 0$. make this selection so that I_x depends measurably on x.

Set

$$E^m = \{x \in E : |I_x| = 2^m\beta\}, \tag{36}$$

for $0 \le m \le C\log(1/\beta)$. Partition a compact subset of $\mathbb{R}$ into $\sim (2^m\beta)^{-1}$ intervals J_j^m of lengths $2^m\beta$, and set

$$\begin{aligned} E_j^m &= \{x \in E^m : I_x \cap J_j^m \neq \emptyset\} \\ F_j^m &= F \cap \pi^{-1}(J_{j-1}^m \cup J_j^m \cup J_{j+1}^m). \end{aligned} \tag{37}$$

The collections $\{E_j^m\}_{m,j}$ and $\{\pi(F_j^m)\}_j$ have bounded overlaps, in the sense that there exists an absolute constant $C < \infty$ such that no point of E belongs to more than C sets E_j^m, and for fixed m, no point of $\pi(F)$ belongs to more than C sets F_j^m. Since $E^m = \cup_j E_j^m$, (35) and (34) imply that

$$\mathcal{X}(E^m, F) \le \sum_j \mathcal{X}(E_j^m, F) \le 2\beta \sum_j |E_j^m| \le C2^{\varepsilon m} \sum_j \mathcal{X}(E_j^m, F_j^m). \tag{38}$$

To each pair F_j^m, E_j^m we associate the two quantities

$$\alpha_j^m = \mathcal{X}(E_j^m, F_j^m)/|F_j^m| \qquad \text{and} \qquad \beta_j^m = \mathcal{X}(E_j^m, F_j^m)/|E_j^m|. \tag{39}$$

Now (34) guarantees that $\beta_j^m \ge c2^{-\varepsilon m}\beta$, and $\beta_j^m \le 2\beta$ since

$$\beta_j^m = |E_j^m|^{-1} \int_{E_j^m} X^*(\chi_{F_j^m}) \le \sup_E X^*(\chi_F) \le 2\beta.$$

Fix m, j. By the construction summarized in Lemma 1.4, there exist sets $\Omega_k^{m,j}$, defined for each $1 \le k \le d$, satisfying the conclusions of that lemma with α_j^m, β_j^m respectively replacing α, β. Moreover, as a consequence of (34), these may be constructed so as to enjoy an additional property: for each even k, for each $\omega \in \Omega_k^{m,j}$, for any subinterval I of $\mathbb{R}$ having length $2^{m-1}\beta$,

$$|\{t \notin I : (\omega, t) \in \Omega_{k+1}^{m,j}\}| \ge c2^{-\varepsilon m}\beta \ . \tag{40}$$

2.2.7. *This is now routine.* To shorten the discussion I will henceforth assume that the dimension d is even; in the odd-dimensional case all of the formulas, but none of the conceptual aspects, are changed. See [11] for the necessary adjustments.

Arguing as in §1 of these notes, but exploiting the fact that for any indices $\mu < \nu$ and any t_μ, $|t_\nu - t_\mu| \gtrsim 2^m\beta$ for a set of values of t_ν having measure $\gtrsim 2^{-\varepsilon m}\beta$, one can deduce that if $\varepsilon > 0$ is chosen to be sufficiently small then

$$\int_{\Omega_d^{m,j}} |J|\, ds\, dt \ge c2^{3m/2}(\alpha_j^m)^d \cdot (\beta_j^m)^{d(d-1)/2}, \tag{41}$$

from which it follows that

$$|E_j^m| \ge c2^{m/2}(\alpha_j^m)^d(\beta_j^m)^{(d^2-d-2)/2}(2^m\beta).$$

This may be reexpressed as

$$\mathcal{X}(E_j^m, F_j^m) \le C2^{-\delta m}|E_j^m|^{d/(d+2)}|F_j^m|^{2d/(d+2)(d-1)} \cdot (2^m\beta)^{-2/(d+2)(d-1)} \tag{42}$$

for some positive constants $C < \infty,\ \delta > 0$.

The loss of a bounded number of factors of $2^{-m\varepsilon}$ has been more than adequately compensated for by the superlinear growth of the Jacobian factor away from the "diagonal".

2.2.8. *How $|\pi(F)|$ enters the estimates.* It remains to turn (42) into a mixed-norm estimate. The projection $\pi(F_j^m)$ of $F_j^m \subset \mathbb{R}^1 \times \mathbb{R}^{d-1}$ onto the first factor, $\mathbb{R}^1$, satisfies $|\pi(F_j^m)| \le 3 \cdot 2^m\beta$, simply because $\pi(F_j^m) \subset J_{j-1}^m \cup J_j^m \cup J_{j+1}^m$ and each of the latter intervals has length $2^m\beta$. Substituting this into (42), we obtain

$$\mathcal{X}(E_j^m, F_j^m) \le C2^{-\delta m}|E_j^m|^{d/(d+2)}\big(|F_j^m| \cdot |\pi(F_j^m)|^{-1/d}\big)^{2d/(d+2)(d-1)}.$$

Invoking Lemma 2.2, with $q = q_0, r = r_0$, we find that

$$\mathcal{X}(E_j^m, F_j^m) \le C2^{-\delta m}|E_j^m|^{d/(d+2)}\|\chi_{F_j^m}\|_{L^{q'}(L^{r'})}.$$

2.2.9. *Conclusion of proof.* To conclude the proof it suffices to sum over j, invoking Hölder's inequality and exploiting the bounded overlap property of the projections $\pi(F_j^m)$ (which holds uniformly in m), and then to sum over m, exploiting the factor of $2^{-\delta m}$. $\square$

2.3. **The two-dimensional case.** The argument, as given, fails to apply[7] in the two-dimensional case, where one is dealing with the Kakeya maximal function and the exponent r_0 equals ∞. It may be of interest to see how our argument can be reorganized so as to apply in dimension two. We begin with a point $(\theta_0, y_0) \in F$ rather than with a point in E, and compute $\gamma^*(\gamma(\theta_0, y_0, s), t) = (t, y_0 + s(\theta_0 - t))$. Thus $|J(s,t)| = |t - \theta_0|$ and we immediately deduce that $|F| \ge c\beta^2\alpha$. More precisely, introducing sets E_j^m, F_j^m as above, we find that

$$|F_j^m| \ge c\alpha_j^m\beta_j^m \cdot 2^{m(1-\varepsilon)}\beta \ . \tag{43}$$

[7]The difficulty is that no factors of $t_i - t_j$ arise for $d = 2$, so no power of $2^m\beta$ appears in the lower bound for $\int_{\Omega_d^{m,j}} |J|$, so $|\pi(F_j^m)|$ can't be brought into play in the same way as in higher dimenions.

Since $|\pi(F_j^m)| \le c2^m\beta$, this implies that

$$|F_j^m| \ge c\frac{\mathcal{X}(E_j^m,F_j^m)}{|F_j^m|}\cdot\frac{\mathcal{X}(E_j^m,F_j^m)}{|E_j^m|}\cdot|\pi(F_j^m)|\,\beta^{C\varepsilon}, \tag{44}$$

which leads to

$$\mathcal{X}(E_j^m,F_j^m) \le C|E_j^m|^{1/2}|F_j^m||\pi(F_j^m)|^{-1/2}\beta^{-C\varepsilon}. \tag{45}$$

Since $|F_j^m| \le |\pi(F_j^m)|^{1/2}\|\chi_{F_j^m}\|_{L^2(L^1)}$ by (32), we conclude that

$$\mathcal{X}(E_j^m,F_j^m) \le C|E_j^m|^{1/2}|\chi_{F_j^m}|_{L^2(L^1)}\beta^{-C\varepsilon}\ . \tag{46}$$

Summing over j,k and applying Cauchy-Schwarz as in the case $d \ge 3$ gives

$$\mathcal{X}(E,F) \le C|E|^{1/2}\|\chi_F\|_{L^2(L^1)}\beta^{-C\varepsilon}, \tag{47}$$

whence $\mathcal{X}(E,F) \le C|E|^{1/2}\|\chi_F\|_{L^2(L^r)}$ for all $r > 1$. Thus (by virtue of a small interpolation argument) X maps L^2 boundedly to $L^2(L^r)$ for all $r < \infty$. □

2.3.1. *Exercise.* Use the ideas introduced in §2 to reprove Theorem 1.2 using only Ω_d and avoiding the band structure argument.

2.3.2. *An example.* This analysis raises a general problem: give sharp lower bounds for integrals $\int_\Omega |u(x)|\,dx$, where Ω is an arbitrary parameter space tower (with two parameters α,β), and u is a smooth function. The bound should be phrased in terms of α,β, under the hypothesis that certain partial derivatives $|\partial^\gamma u/\partial x^\gamma|$ are bounded away from zero, and all partial derivatives up to some (higher) order are bounded.[8]

Consider a simple instance of this problem: $\Omega \subset \mathbb{R}^2$; $\Omega \subset [0,1]^2$ with coordinates (x,y), x ranges over a subset of $[0,1]$ having measure β, and for each x, the set of $y \in [0,1]$ for which $(x,y) \in \Omega$ has measure $\sim \alpha$. Suppose that $\partial^2 u/\partial x\partial y$ is bounded below on $[0,1]^2$, and that u is bounded above in C^N norm. Moreover, suppose that $\beta \lesssim \alpha$. Then[9] restriction $\beta \lesssim \alpha$, a natural guess is that $\big|\int_\Omega |u|\big| \gtrsim \alpha^2\beta^2$; this bound is realized when Ω is a rectangle of dimensions (β,α) and $u(x,y) = xy$.

However, this bound is not correct for arbitrary $u \in C^\infty$. The building blocks on which the construction is based are the functions $u(x,y) = y(x - \beta\ln(y/\alpha))$, defined for $10\alpha \le y \le 1$ and $x \in [0,1]$. We assume always that $\beta \le \alpha \ll 1$. Then in the specified region of values of (x,y), these functions u are C^2, uniformly in α,β; moreover $u \in C^N$, uniformly, provided that $\beta \le C\alpha^{N+1}$. And of course $\partial^2 u/\partial x\partial y \equiv 1$.

Let Ω be the set of all points in the specified region where $|x - \beta\ln(y/\alpha)| \le \alpha\beta/y$. Then in that region, $|\partial u/\partial y| = |(x - \beta\ln(y/\alpha)) - \beta| \le \alpha\beta/y + \beta \le 2\beta$. Therefore for each x, the set of all y for which (x,y) belongs to Ω has measure at least $c\alpha$ for some fixed constant $c > 0$, independent of α,β,x. Moreover x varies from approximately $\beta\ln(10)$ to $\beta\ln(1/\alpha)$. Thus Ω is a parameter space tower, with parameters $(\tilde\beta,\alpha) = (\beta\ln(1/\alpha),\alpha)$. But since $u \le \alpha\beta$ on Ω,

$$\int_\Omega |u| \le C\alpha\beta|\Omega| \sim \tilde\beta^2\alpha^2/\ln(1/\alpha).$$

[8] A rather different problem will be discussed in the ICMS lectures of Carbery, in which a lower bound is imposed on $|\Omega|$ (which may be viewed as a parameter space tower of "height" 1, corresponding to a single iteration, in some $\mathbb{R}^d$), and no *upper* bound is imposed on any derivative.

[9] When $\beta \lesssim \alpha$, one can't do any better than $c\beta\alpha^3$ for a lower bound, as the example $u(x,y) = -(y-x)^2$, with Ω chosen to be a suitable tubular neighborhood of the diagonal, demonstrates.

Thus the seemingly natural *a priori* inequality $\int_\Omega |u| \ge c\tilde\beta^2\alpha^2$ is not valid uniformly in $\tilde\beta, \alpha, u$, even with a fixed upper bound on a C^N norm of u, no matter how large N is.

A single C^∞ counterexample u may be constructed by patching these functions together; the details are left to the reader.

3. The simplest multilinear operators

3.1. **Some examples.**

3.1.1. *Standard linear operators.* Some of the most fundamental linear operators arising in Euclidean harmonic analysis are:

- $f \mapsto \hat f(\xi) = \int f(x)e^{-i\xi x}\,dx$ (Fourier transform)
- $Tf(x) = \int_{\mathbb{R}^d} f(y)K(x-y)\,dy$ where $K(z) = |z|^{-d}\Omega(z/|z|)$ and $\Omega \in C^\infty(S^{d-1})$ has mean value zero (homogeneous convolution Calderón-Zygmund singular integral operators)
- $Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)\,dy$ (nonconvolution Calderón-Zygmund operators) where $|K(x,y)| \lesssim |x-y|^{-d}$, $|\nabla_{x,y}K(x,y)| \lesssim |x-y|^{-d-1}$, and K satisfies certain subtle cancellation conditions.
- $I_\alpha f(x) = \int_{\mathbb{R}^d} |x-y|^{-d+\alpha} f(y)\,dy$ where $0 < \alpha < d$ (fractional integral operators)
- $Tf(x) = \int_{\mathbb{R}^d} e^{ix\xi} m(\xi)\hat f(\xi)\,d\xi$ for various classes of multipliers m (Fourier multiplier operators) including the cases where m is
  1. the characteristic function of a half-space,
  2. the characteristic function of a ball,
  3. surface measure on the unit sphere S^{d-1},
  4. a power of $|\xi|$ times surface measure on the cone $|\xi_d| = |(\xi_1,\dots,\xi_{d-1})|$,
  5. a power of $|\xi|$ times surface measure on the parabola $|\xi_d| = |(\xi_1,\dots,\xi_{d-1})|^2$.
- $f \mapsto f * \sigma$ where σ is surface measure on the unit sphere in $\mathbb{R}^d$

Certain related operators which are not quite linear are equally fundamental:

- $Mf(x) = \sup_{r>0} r^{-d}\int_{|x-y|<r} |f(y)|\,dy$ (Hardy-Littlewood maximal operator in $\mathbb{R}^d$)
- $\mathcal{C}^* f(x) = \sup_{\xi\in\mathbb{R}} \left|\int_{\mathbb{R}} f(x-y)e^{i\xi y}\,y^{-1}\,dy\right|$ (the maximal partial sum operator for the Fourier transform)

3.1.2. *Classical multilinear operators.* Fundamental multilinear operators arising in Euclidean harmonic analysis include:

- $(f,g) \mapsto f\cdot g$ (multiplication)
- $(f,g) \mapsto f * g(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy$ (convolution)
- $\pi(f,g)(x) = (2\pi)^{-d}\int_{\mathbb{R}^{d+d}} \chi_{|\xi|\le|\eta|} e^{ix(\xi+\eta)}\hat f(\xi)\hat g(\eta)\,d\xi\,d\eta$ (paraproducts)
- $T(f_1,\dots,f_k)(x) = \int_{\mathbb{R}^{kd}} m(\xi_1,\dots,\xi_d)e^{ix(\xi_1+\dots+\xi_k)}\prod_{j=1}^k \hat f_j(\xi_j)\,d\xi_j$ where $m(r\xi) \equiv m(\xi)$ for all $r \in \mathbb{R}^+$ and $\xi \in \mathbb{R}^{kd}$, and $m \in C^\infty(\mathbb{R}^{kd}\backslash\{0\})$ (multilinear Fourier multiplier operators of Calderón-Zygmund/Coifman-Meyer type)

3.1.3. *More singular multilinear operators.* Various more singular multilinear operators have been receiving increasing attention:

- In $\mathbb{R}^1$, $(f,g) \mapsto \int_{\mathbb{R}} f(x-t)g(x+t)\,t^{-1}\,dt$ (the bilinear Hilbert transform, studied by Lacey and Thiele)

- $$(f_1,\dots,f_k)\mapsto \iint_{x_1<x_2<\cdots<d_k} e^{i\xi(x_1-x_2+x_3-x_4+\cdots\pm x_k)}\prod_{j=1}^k f_j(x_j)\,dx_j \tag{48}$$

  and certain maximal variants analogous to $\mathcal{C}^*$ (multilinear variants of the Fourier transform, analyzed in a rudimentary way by Christ and Kiselev in connection with the spectral theory of Schrödinger operators)
- $I_\alpha(f_1,\dots,f_k)(x)=\int_{\mathbb{R}^d}\prod_{j=1}^k f_j(\ell_j(x,t))|t|^{\alpha-d}\,dt$ where $x\in\mathbb{R}^d$, each function $\ell_j:\mathbb{R}^{d+d}\mapsto\mathbb{R}^d$ is linear, $0<\alpha<d$, and the collection $\{\ell_j\}$ satisfies an appropriate nondegeneracy condition (multilinear generalizations of fractional integrals, investigated by Grafakos and Kalton, and by Kenig and Stein)
- Generalizations of the bilinear Hilbert transform, in which the two functions f,g are replaced by $k\ge 2$ functions, the dimension d of the ambient space is arbitrary, and the integration is taken over an n-dimensional subspace of $\mathbb{R}^{kd}$, with $t^{-1}\,dt$ replaced by an n-dimensional Calderón-Zygmund-type convolution kernel (investigated by Muscalu, Tao, and Thiele).
- The bilinear ball multiplier $T(f,g)(x)=\int_{\mathbb{R}^{2d}}e^{ix(\xi+\eta)}\hat f(\xi)\hat g(\eta)m(\xi,\eta)\,d\xi\,d\eta$ where $x\in\mathbb{R}^1$, f,g are likewise defined on $\mathbb{R}^1$, and m is the characteristic function of the unit ball $\xi^2+\eta^2<1$ in $\mathbb{R}^2$ (studied by Grafakos and Li).

These matters remain only partially understood. Among the fundamental unresolved questions are:

- Are there any L^p estimates at all for *trilinear* analogues $\int f(x+t)g(x-t)h(x-\theta t)\,t^{-1}\,dt$ of the bilinear Hilbert transform (where $\theta\ne\pm1$)? (Although the theory of Muscalu, Tao, and Thiele applies to certain trilinear operators, this one is too singular.)
- Do certain fully nonlinear variants of the Fourier transform (48), which arise in Schrödinger operator scattering theory [13], act on L^2? (They have been shown to act on L^p whenever $p\in[1,2)$.)

3.2. **The simplest multilinear operators of all.**

3.2.1. *Introduction.* In these notes I discuss some of the simplest of all multilinear operators:

$$T(f_1,\dots,f_m)(x)=\int_{|t|\le 1}\prod_{j=1}^m f_j(S_j(x,t))\,dt. \tag{49}$$

Here the variables x,t both belong to $\mathbb{R}^d$, and the $S_j:\mathbb{R}^{d+d}\mapsto\mathbb{R}^d$ are surjective linear mappings. Writing $S_0(x,t)\equiv x$, we exclude degenerate cases by assuming always that for any two indices $i\ne j$, the mapping $(x,t)\mapsto(S_i(x,t),S_j(x,t))$ is invertible; in that case we say that $\{S_j:0\le j\le 3\}$ is nondegenerate.

A concrete example, with dimension $d=1$ and degree $m=3$, is

$$T_\theta(f,g,h)(x)=\int_{|t|<1} f(x+t)g(x-t)h(x-\theta t)\,dt \tag{50}$$

for any $\theta\notin\{-1,1\}$. These are far less singular cousins of the trilinear Hilbert transform. The variant

$$T_\infty(f,g,h)(x)=\int_{|t|<1} f(x+t)g(x-t)h(t)\,dt \tag{51}$$

is likewise included in the framework.

The cases $m = 1, 2$ are well understood. The former amounts to convolution of a function with the characteristic function of the unit ball; under the nondegeneracy hypothesis, linear changes of variables reduce the latter to a localized form of convolution, $(f_1, f_2) \mapsto f_1 * f_2$.

This operation is benign in three respects: (i) There is no singular kernel; the integration is with respect to Lebesgue measure on $\mathbb{R}^d$. (ii) The restriction $|t| \le 1$ likewise ensures that any L^p estimates reduce to a bounded subregion of $\mathbb{R}^d$. (iii) There is no (overt) oscillation/cancellation in the problem; $|T(f_1, \cdots, f_m)| \le T(|f_1|, \dots, |f_m|)$. We therefore may, and will, assume henceforth that all functions are nonnegative, and (by point (ii)) that they are all supported in a fixed compact subset of $\mathbb{R}^d$.

Nonetheless, there is a certain kind of singularity here. If one seeks an L^1 estimate, one finds that $\int_{|x|\le 1} T(f_1, \dots, f_m)(x)\,dx$ is essentially $\int_{V\subset\mathbb{R}^{md}} \prod_j f_j(y_j)\,d\sigma(y_1, \dots, y_m)$ where $V \subset \mathbb{R}^{md}$ is a linear subspace of dimension $2d$, and σ is Lebesgue measure on V. For any $m > 2$, V is a proper subspace of $\mathbb{R}^{md}$. This formula expresses the fact that for $m \ge 3$, for *generic* points $(x_1, \dots, x_m) \in \mathbb{R}^{md}$, the value of $(f_1, \dots, f_m)$ at $x = (x_1, \dots, x_m)$ makes no contribution to $T(f_1, \dots, f_m)$; conversely, points $x \in V$ make disproportionately large contributions.

We will see that the theory of L^p estimates for these operators is subtle, for a certain range of p, even in the simplest nontrivial case $m = 3$, $d = 1$. Connections with the Kakeya problem — or at least, with recent analysis of that problem — will also emerge.

3.2.2. *The simplest question.* The simplest question about L^p estimates is: For which exponents p is such an operator T *defined* on $L^p \times L^p \times \cdots \times L^p$? That is, for which p is $T(f_1, \dots, f_j)(x)$ finite for almost every $x \in \mathbb{R}^d$, for arbitrary nonnegative functions $f_j \in L^p$?

3.3. **Routine observations and preliminaries.**

3.3.1. *The obvious estimates.* Let T be an operator of the class (49), satisfying the nondegeneracy hypothesis. Whenever T is defined on some L^p, it is defined on L^r for any $r > p$. This holds because the restriction of $T(f_1, \dots, f_m)$ to any particular bounded region depends only on the restrictions of $f_1, \dots, f_m$ to a certain larger bounded region.

In the bilinear case, one has

$$\int T(f_1, f_2)(x)\,dx = c \iint f_1(y_1) f_2(y_2)\,dy_1\,dy_2$$

by a linear change of variables (whose justification relies on the nondegeneracy hypothesis), so $\|T(f_1, f_2)\|_{L^1} \le c\|f_1\|_{L^1}\|f_2\|_{L^1}$, with equality for nonnegative functions. As a corollary one has for any m

$$\|T(f_1, f_2, 1, \dots, 1)\|_{L^1} \le C\|f_1\|_{L^1}\|f_2\|_{L^1}$$

where 1 denotes the constant function $\equiv 1$. The hypotheses are of course symmetric in all the f_j, so this bound can be restated as

$$\|T(f_1, f_2, \dots, f_m)\|_{L^1} \le C\|f_i\|_{L^1}\|f_j\|_{L^1} \prod_{k\ne i,j} \|f_k\|_{L^\infty}$$

for any $i \ne j \in \{1, 2, \dots, m\}$. A simple multilinear interpolation (see for instance [25]) then gives

$$\|T(f_1, \dots, f_m)\|_{L^1} \le C \prod_{j=1}^{m} \|f_j\|_{L^p} \qquad \text{where } p = m/2. \tag{52}$$

An equally simple estimate is that for any index j,

$$\||T(f_1, \ldots, f_m)\|_{L^\infty} \le C\|f_j\|_1 \prod_{k\ne j} \|f_j\|_\infty.$$

Interpolation then yields the general bound

$$\||T(f_1, \ldots, f_m)\|_{L^q} \le C \prod_{j=1}^{m} \|f_j\|_p \ \text{ where } 1 + q^{-1} = mp^{-1},$$

whenever $1 \le q \le \infty$. We have thus proved that T is defined on $L^p \times \cdots L^p$ for all $p \ge m/2$.

A necessary condition for T to be bounded from $(L^p)^m$ to L^q (by which we mean L^q of a fixed bounded ball in $\mathbb{R}^d$) can be derived as follows. Take $f_1 = f_2 = \cdots = f_m$ each to be the characteristic function of a ball of small radius δ centered at 0. Then $T(f_1, \ldots, f_m)(x) \ge c\delta^d$ whenever $|x| \le c\delta$, for some fixed constant c. Thus for any $0 < q < \infty$, $\|T(f_1, \ldots, f_m)\|_q \sim \delta^d \delta^{d/q}$, while $\prod_{j=1}^m \|f_j\|_p = C\delta^{md/p}$. By letting $\delta \to 0$ we find that a necessary condition is that $1 + q^{-1} \ge mp^{-1}$.

3.3.2. *The simplest question, refined.* We may thus refine our question: Does there exists $p < m/2$ for which T is defined on $(L^p)^m$? Of course, various refinements come immediately to mind.

It is natural to expect that if $T(f_1, \ldots, f_m)$ is finite almost everywhere for all $f \in (L^p)^m$, then T actually maps $(L^p)^m$ boundedly to L^q for some q. We have shown that q must satisfy $1 + q^{-1} \ge mp^{-1}$; for $p < m/2$ this forces $q < 1$.

3.3.3. *Rearrangement.* Denote by f^* the symmetric decreasing rearrangement of a nonnegative function f in $\mathbb{R}^d$. The function $f^* : \mathbb{R}^d \mapsto [0, +\infty)$ is radially symmetric and nonincreasing along radii extending from the origin, it is equimeasurable with $|f|$ in the sense that for any $\alpha > 0$, $|\{x : f^*(x) > \alpha\}| = |\{x : |f(x)| > \alpha\}|$, and it is lower semicontinuous.

Then for any of our operators T,

$$\|T(f_1, \ldots, f_m)\|_q \le \|T(f_1^*, \ldots, f_m^*)\|_q \ \text{ for any } q \ge 1. \tag{53}$$

See Theorem 3.8 in [31], the original paper [5], or the ICMS lectures of Barthe. In particular, for $q \ge 1$, if the functions f_j are all taken to be characteristic functions of sets, and if the measures of those sets are prescribed, then $T(\chi_{E_1}, \ldots, \chi_{E_m})$ is maximized in L^q norm if each set E_j is a ball centered at the origin.[10]

On the other hand, if each f_j is the characteristic function of a ball $B(0, \delta_j)$ centered at the origin, then it can be checked that $\|T(f_1, \ldots)\|_q \le C \prod_j \|f_j\|_p$, whenever $1 + q^{-1} = mp^{-1}$, for all $q > 0$.

However, for $q < 1$ this rearrangement inequality is false. Indeed, we will see that the extremals for the L^q norm inequality can then have a different character altogether.

[10] It is interesting that the more singular operators appearing on the right-hand side of (2) also satisfy such a rearrangement inequality, even though they are not covered by the Brascamp-Lieb-Luttinger rearrangement inequality (53). See [7].

3.3.4. *Conservation law.* There is a kind of L^1 conservation law: For nonnegative functions f_1, f_2 supported in a sufficiently small neighborhood of the origin (so that the restriction $|t| < 1$ in the integral defining T becomes moot),

$$\|T(f_1, f_2, 1, \cdots, 1)\|_1 = c\|f_1\|_1\|f_2\|_1 \tag{54}$$

for a certain constant $c > 0$.

In proving that T maps $(L^p)^m$ to some L^q, it suffices to treat the case where each f_j is the characteristic function of a set; the method of real interpolation continues to apply when the target exponent q is < 1. See [21] for a recent discussion of multilinear interpolation with exponents < 1 and with weak type estimates, relying on the ideas introduced by Wolff [57] in connection with reiteration.

Consider for simplicity the case where $f_j = \chi_{E_j}$ where all the sets E_j have measures comparable to an arbitrary small $\delta > 0$. Then the conservation law gives

$$\|T(\chi_{E_1}, \chi_{E_2}, \cdots, \chi_{E_m})\|_1 \le \|T(\chi_{E_1}, \chi_{E_2}, 1, \cdots, 1)\|_1 = c|E_1| \cdot |E_2| \sim \delta^2.$$

On the other hand, $\prod_{j=1}^m \|f_j\|_p = \prod_{j=1}^m |E_j|^{1/p} \sim \delta^{m/p}$. Thus for $p = m/2$ and $q = 1$, $\|T(\chi_{E_1}, \chi_{E_2}, 1, \cdots, 1)\|_1 \le \prod_{j=1}^m \|f_j\|_p$. Consequently $T(\chi_{E_1}, \chi_{E_2}, \cdots, \chi_{E_m})$ will be much smaller in L^1 norm than $\prod_j |E_j|^{1/p}$, *unless* it is nearly as large in L^1 norm as $T(\chi_{E_1}, \chi_{E_2}, 1, \cdots, 1)$.

For small δ, this is not the case for typical sets. If each E_j is the characteristic function of a ball centered at $x_j \in \mathbb{R}^d$ of radius δ, and if $m > 2$, then $T(\chi_{E_1}, \chi_{E_2}, \cdots, \chi_{E_m})$ will vanish identically, with very high probability, if δ is sufficiently small and the centers x_j of the balls are chosen independently at random (with respect to Lebesgue measure) in the unit ball of $\mathbb{R}^d$. This holds because $\int T(\chi_{E_1}, \chi_{E_2}, \cdots, \chi_{E_m})(x)\,dx$ equals a constant times the integral of $\prod_j \chi_{E_j}$ over a $2d$–dimensional subspace of $\mathbb{R}^{md}$; random δ–balls will miss this subspace with high probability.

We have already noted that T does behave well in the "maximally coincident" case, when all the balls are centered at the same point. This, together with the observation of the preceding paragraph, might suggest that T is likely to be defined on L^p for (at least some) $p < m/2$.

3.3.5. *Phase transition.* This problem exhibits a sort of phase transition at $p = m/2$, $q = 1$. In L^1 there is a kind of conservation law, as we have seen. Consider a function having prescribed L^1 norm. For $q > 1$, its L^q norm becomes large when it is concentrated, that is, when much of its L^1 mass is supported on a set of small measure. For $q < 1$, however, the situation is reversed; the L^q norm becomes larger when the function is diffuse, in the sense that (a substantial fraction of) its L^1 mass is spread out somewhat evenly over a set having large measure.

Our problem, for $q < 1$, seems to be governed by the interplay between two conflicting considerations. On one hand, the L^q norm is larger when there are few coincidences, in the sense that $T(f)$ is spread out over a comparatively large set. On the other hand, $\|T(f_1, \ldots, f_m)\|_1$ will be much smaller than $\prod_j \|f_j\|_p$ unless there are many coincidences, in the sense that the events that $f_j(S_j(x,t))$ is comparatively large for $j = 1, \ldots, m$ are strongly correlated. And when $\|T(f_1, \ldots, f_m)\|_1$ is much smaller than $\prod_j \|f_j\|_p$, it is likely that $\|T(f_1, \ldots, f_m)\|_1$ will be dominated by $C\prod_j \|f_j\|_r$ for some $r < p$.

3.4. **Statements of results.** The main theorems apply only to the simplest nontrivial case of trilinear operators, $m = 3$. The threshold index is then $p = 3/2$. For simplicity I will restrict attention to dimension $d = 1$, although the results do extend to higher dimensions.

Theorem 3.1. *Let $m = 3$ and $d = 1$. Suppose that theta $\notin \{\pm 1\}$. If $\theta \in \mathbb{Q}$ or if $\theta = \infty$, then there exists $p < 3/2$ such that T_θ maps $(L^p)^3$ boundedly to L^q, where $1 + q^{-1} = 3p^{-1}$.*

If on the other hand $\theta \notin \mathbb{Q}$, then for any $p < 3/2$, there exist nonnegative functions $f_j \in L^p$ and a set $E \subset \mathbb{R}^d$ of positive Lebesgue measure such that $T_\theta(f_1, f_2, f_3)(x) = +\infty$ for all $x \in E$.

Given any $p < 3/2$, there exist $\theta \in \mathbb{Q}$, nonnegative functions $f_j \in L^p$, and a set E as above, such that for all $x \in E$, $T_\theta(f_1, f_2, f_3)(x) = +\infty$.

For any $\theta \in \mathbb{R} \cup \{\infty\}$ there exist $p > 1$, functions $f_j \in L^p$, and a set E as above, such that $T_\theta(f_1, f_2, f_3)(x) = +\infty$ for all $x \in E$.

Thus for each $\theta \in \mathbb{Q} \cup \{\infty\}$ there exists $p_0(\theta) \in (0, 3/2)$ such that T_θ is defined on $(L^p)^3$ for all $p > p_0$ and for no $p < p_0$. There is no θ for which I know the exact value of $p_0(\theta)$.

In dimensions $d \geq 2$, the necessary and sufficient on the linear mappings S_j for T to be defined on $(L^p)^3$ for some $p < 3/2$ is *rational commensurability*, defined as follows:

Definition. *Suppose that $\{S_j : 0 \leq j \leq 3\}$ is nondegenerate. The collection of linear mappings $\{S_j\}$ is said to be rationally commensurate if there exist linear automorphisms h_j of $\mathbb{R}^d$ such that the vector subspace of endomorphisms of $\mathbb{R}^2$ generated by $\mathbb{Q}$-linear combinations of $\{h_j \circ S_j : 0 \leq j \leq 3\}$ has dimension two over $\mathbb{Q}$.*

Example. Rational commensurability is not a matter of rationality of matrix entries: In $\mathbb{R}^2$, the operator

$$\int f_1(x+t) f_2(x-t) f_3(x - \Theta t)\, dt \qquad \text{with} \qquad \Theta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is *not* rationally commensurate.

3.4.1. *Remarks on $m = 4$.* No necessary and sufficient condition is known (to this author) for $m \geq 4$.

- Consider the case $m = 4$ in dimension $d = 1$:

$$T_4(f_1, \cdots, f_4)(x) = \int \prod_{j=1}^{4} f_j(x - \theta_j t)\, dt$$

  with $\theta_1 = 1, \theta_2 = -1$. The construction used in the trilinear case for irrational θ yields a negative result for $m = 4$ when θ_3, θ_4 are simultaneously well approximable by rationals, in the sense that there exist two sequences of rationals $r_{3,\nu}/q_\nu$ and $r_{4,\nu}/q_\nu$ with common denominators q_ν tending to ∞ such that $\left|\theta_j - \frac{r_{j,\nu}}{q_\nu}\right| \leq \frac{C}{q_\nu^2}$ for $j = 3, 4$.
- It is a direct consequence of the trilinear theorem that T is defined on $(L^p)^4$ for some $p < m/2 = 2$ when at least one of θ_3, θ_4 is rational.
- Left unresolved is the case where θ_3, θ_4 are both irrational and are not well approximable by rationals with the same denominators.

3.5. **Connection with additive number theory.**

3.5.1. *Discrete analogues.* Denote by $\#(S)$ the cardinality[11] of a finite set S.

[11]Sorry: Sometimes I write $\#(S)$, sometimes $|S|$.

Theorem 3.2. *Suppose that (m,k), $(m',k') \in \mathbb{Z}^2$ and that no two of the vectors $(1,0)$, $(0,1)$, (m,k), (m',k') are linearly dependent. Then there exists $p < 3/2$, depending only on m,k,m',k', such that for any torsion-free Abelian group G, for any finite subsets $A,B,C \subset G$, the multiplicity function*

$$\mu(x) = \#(\{(a,b,c) \in A \times B \times C : c = m'a + k'b \text{ and } x = ma + kb\})$$

satisfies

$$\sum_{x \in G} \mu(x)^q \leq \#(A)^{q/p} \#(B)^{q/p} \#(C)^{q/p}$$

where $1 + q^{-1} = 3p^{-1}$.

Remarks.

- Again, $\|\mu\|_{\ell^q} \leq [|A| \cdot |B| \cdot |C|]^{1/p}$ whenever $1 + q^{-1} = 3^{-1}$ and $3/2 \leq p \leq 3$, by the same reasoning as in the continuum case.
- Such a universal result can only hold for $1 + q^{-1} \geq 3p^{-1}$.
- This theorem is actually equivalent to the positive part of Theorem 3.1 for $\theta \in \mathbb{Q} \cup \{\infty\}$; either can be deduced from the other by appropriate limiting arguments. Alternatively, both follow from the same proof.
- In contrast to the local continuum case treated in Theorem 3.1, here the L^q norm becomes stronger as q decreases.

Corollary 3.3. *Let $(m,k), (m',k')$ satisfy the hypotheses of Theorem 3.2. There exists $\varepsilon > 0$, depending only on m,k,m',k', such that for any torsion-free Abelian group G, any positive integer N, and any finite subsets $A,B,C \subset G$ such that $\#(A), \#(B), \#(C) \leq N$,*

$$\#(\{ma + kb : (a,b,m'a + k'b) \in A \times B \times C\}) \leq N^{2-\varepsilon}.$$

This corollary follows directly from the theorem since if $x = ma + kb$ with $(a,b,m'a + k'b) \in A \times B \times C$, then $\mu(x) \geq 1$.

The case where $(m,k) = (1,-1)$ and $(m',k') = (1,1)$ was treated by Bourgain [3], and an improved estimate for ε was obtained by Katz and Tao [27], by a different argument. See also Gowers [19], Proposition 1.2.

3.5.2. *Freiman's theorem (as extended by Ruzsa).* For any two subsets A, B of an Abelian group G, denote by $A + B$ the sum-set $\{a + b : (a,b) \in A \times B\}$.

Ruzsa's extension [41] of Freiman's theorem asserts the following. If A, B are finite subsets of a torsion-free Abelian group with $|A| \sim |B|$, and if

$$|A + B| \leq K|A|,$$

then A has a kind of arithmetic structure: $A \subset P$ for some arithmetic multiprogression P of dimension bounded by $C(K)$ and of cardinality $\leq C'(K)|A|$.

Definition. *An arithmetic multiprogression in an Abelian group G is a set of the form $\{a + n_1 q_1 + n_2 q_2 + \cdots n_r q_r : 1 \leq n_j \leq k_j\}$ for some positive integers $r, k_1, \ldots, k_r$ and elements $a, q_1, \ldots, q_r \in G$. Bounded dimension means bounded r.*

A related result is the Balog-Szemerédi theorem: Let there be given three subsets A, B, C of an Abelian group, and $R \subset A \times B \times C$ such that for every $(a,b,c) \in R$, $c = a + b$. Suppose that A, B, C have cardinalities $\leq N$, and that $|R| \geq \delta N^2$. Then there exists a subset $\tilde{A} \subset A$ with $|\tilde{A}| \geq cN$ such that $|\tilde{A} + \tilde{A}| \leq CN$. Here c, C depend only on δ. Freiman's theorem then implies that $\tilde{A}$ is a large subset of an arithmetic multiprogression.

Ruzsa [42, 40] established the following bounds, which are weaker in that they give no structural information on A, but which are more directly connected with the type of estimates we seek for multilinear operators: If A, B are subsets of any Abelian group, and if $|A+B| \le K|B|$, then for any positive integers k, l, $|kA-lA| \le K^{k+l}|B|$. Here $kA-lA$ is the set of all sums $\sum_{j=1}^{k} a_j - \sum_{i=1}^{l} a'_i$ with $a_j, a'_i \in A$. In particular, if $B = A$ and $(k, l) = (1, 1)$, then the difference set $|A - A|$ has cardinality[12] $\le C(K)|A|$ if $|A + A| \le K|A|$.

From the perspective of our problem, these theorems apply to an extreme situation, where $|A + B|$ is as small as possible, up to a constant factor. In the discrete version of our multilinear operator problem much less is given; merely a *subset* of $A \times B$ whose projection under the addition map $(a, b) \mapsto a + b$ has cardinality $\lesssim |C|$. The essential complication here is that only a subset of $A \times B$ is constrained. On the other hand, in order to prove Corollary 3.3 as stated, one aims to conclude much less, merely that its projection under subtraction has cardinality $\ll |A|^{2-\varepsilon}$, rather than $\lesssim |A|$. See [19, 3, 27]. Our positive results are based on a further development of the formalism of the latter paper.

As will be seen below, arithmetic progressions do figure prominently in the theory of our multilinear operators.

For more information on these matters, see the ICMS lectures of Green.

3.6. **Negative results.**

3.6.1. *The irrational case.* Consider the case of trilinear operators T_θ in dimension $d = 1$, with irrational θ. We want to show that for any $p < 3/2$, T fails to be defined on $(L^p)^3$. Let $\delta > 0$ be a small parameter.

Let the functions f_1, f_2, f_3 be the characteristic functions of sets A, B, C; to simplify notation I will write $T(A, B, C)$ for $T_\theta(\chi_A, \chi_B, \chi_C)$. The aim is to construct A, B, C with comparable measures $\sim \delta$, so that $\|T(A, B, C)\|_q \sim \|T(A, B, C)\|_1$ no matter how close q is to 0, and moreover so that $\|T(A, B, C)\|_1 \sim \delta^2$. The latter is desirable since $\prod \|f_j\|_p \sim \delta^{3/p}$ is $\ll \delta^2$ whenever $p < 3/2$. The former is essentially necessary; whenever A, B, C have the property that $\|T(A, B, C)\|_q \lesssim \delta^\varepsilon \|T(A, B, C)\|_1$ for some fixed $\varepsilon > 0$, we would have $\|T(A, B, C)\|_q \lesssim \delta^\varepsilon \prod_j \|f_j\|_{3/2} \sim \prod_j \|f_j\|_p$ for some $p(\varepsilon) < 3/2$.

Since we are always working in a fixed bounded region, we'd thus like to make $T(A, B, C)(x) \sim \delta^2$ for a set of values of x having measure ~ 1. That would be the maximum possible L^1 norm, with the maximum possible dispersion, up to bounded factors. For comparison, note that when A, B, C equal the interval $(-\delta, \delta)$, $T(A, B, C)$ is supported on $(-C\delta, C\delta)$ and $T(A, B, C)(x) \sim \delta$ for all x in a subinterval of length comparable to δ.

We will define A, B, C to be *δ–neighborhoods of arithmetic progressions* of cardinality $\delta^{-1/2}$, with spacings comparable to $\delta^{1/2}$. Thus $A \times B$ is a union of $\sim \delta^{-1}$ squares in $\mathbb{R}^2$, and $T(A, B, \mathbb{R})$ thus decomposes as a sum of $\sim \delta^{-1}$ constituents, each of which is supported on an interval of length $\sim \delta$, and has L^∞ norm $\sim \delta$. $T(A, B, C)$ is bounded above by $T(A, B, \mathbb{R})$, but could potentially be much smaller, depending on how often it happens that $(x + t, x - t) \in A \times B$ implies that $x - \theta t \in C$. We aim to construct these sets so that $(x + t, x - t) \in A \times B$ always implies $x - \theta t \in C$, and so that the $\sim \delta^{-1}$ constituents of $T(A, B, C) \equiv T(A, B, \mathbb{R})$ have disjoint supports. Since the support of each such constituent is an interval of length $\sim \delta$, and we are working in a fixed bounded region, there is just enough room to permit this.

[12] More precisely: If $|A| = N$ and $|A + A| = \kappa N$ then $|kA - lA| \le \kappa^{2k+2l} N$.

Since $(1+\theta)/(1-\theta)$ is irrational, there exists [24] a sequence of natural numbers $q_\nu \to \infty$ for which

$$\frac{1+\theta}{1-\theta} = \frac{r_\nu}{q_\nu} + O(q_\nu^{-2}), \quad r_\nu \in \mathbb{Z}.$$

Fix ν, and drop the subcripts to simplify notation. Let $\delta = q_\nu^{-2} = q^{-2}$. Define A, B to be $c_0\delta$–neighborhoods of arithmetic progressions with spacing r^{-1}, q^{-1}, where c_0 is a small constant.

Then for any integers $1 \le i \le r^{-1}$, $1 \le j \le q^{-1}$, there exist $k \in \mathbb{Z}$ such that

$$(1+\theta)\frac{i}{r} + (1-\theta)\frac{j}{q} = \frac{k}{y} + O(\delta)$$

where $y^{-1} = (1-\theta)/q$. Defining C to be a $C_0\delta$–neighborhood of the finite arithmetic progression $\{k/y : 1 \le k \lesssim y\}$, for a sufficiently large constant C_0, the preceding relation means that as in the hypothesis of Freiman's theorem,

$$(1+\theta)A + (1-\theta)B \subset C.$$

In other words, if $(x-t, x+t) \in A \times B$ then $x - \theta t \in C$.

On the other hand, the sum set $A + B$ is enormous; since r, q are relatively prime,

$$\Big|\big(\frac{i}{r} + \frac{j}{q}\big) - \big(\frac{i'}{r} + \frac{j'}{q}\big)\Big| \ge r^{-1}q^{-1} \sim \delta \ \text{ if } (i,j) \ne (i',j').$$

So the $\sim q^2$ "blobs" of which $T_\theta(A,B,C)$ is a sum all have disjoint supports; thus we have very strong "dispersion". In fact the union of their supports has measure ~ 1, uniformly as $q_\nu \to \infty$.

Thus $|A| \sim |B| \sim |C| \sim q^{-1}$ while

$$T_\theta(A,B,C) \sim q^{-2} \text{ on a set of measure} \sim 1. \tag{55}$$

Thus $T(\chi_A, \chi_B, \chi_C) \equiv T(\chi_A, \chi_B, \chi_\mathbb{R})$ and hence $\|T(\cdot)\|_{L^1}$ is not $\ll |A| \cdot |B| \sim (|A| \cdot |B| \cdot |C|)^{2/3}$; the trivial $p = 3/2$ inequality is "saturated" in this example.

(55) expresses dispersion: expanding $T(\cdot)$ in terms of the $\sim q^2$ "blobs" in $A \times B$ gives rise to $\sim q^2$ disjointly supported functions. The union of their supports has measure ~ 1. Hence no matter how small an exponent $s > 0$ may be, the L^s norm is comparable to the L^1 norm. Hence there is no improvement over the trivial $[L^{3/2}]^3 \mapsto L^1$ estimate.

3.6.2. *Remarks.*

- This construction can be carried out for rational θ, but for only finitely many scales; whereas for irrational θ it can be done for an infinite sequence of scales. The universal rate of approximation $|\frac{1+\theta}{1-\theta} - \frac{r_\nu}{q_\nu}| = O(q_\nu^{-2})$ is exactly what is needed for the argument.
- It helps to view this construction geometrically, in terms of projections in differing directions, as illustrated by a slide I'll show during the lecture.
- For rational θ, T_θ makes sense in the discrete setting, after an appropriate change of variables is made to clear denominators. This is not the case for irrational θ. However, this construction can be interpreted in the discrete setting as establishing a lack of *uniformity* of estimates for certain sequences (namely, close rational approximants of irrationals) of rational θ.

3.6.3. *Two more counterexamples.* Two other types of counterexamples are used to establish the other negative conclusions of Theorem 3.1. The two-dimensional example with $\Theta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, indicated above, can be generalized to show that any higher-dimensional trilinear operator either satisfies no estimate below the $L^{3/2}$ threshold, or is essentially one-dimensional. The counterexample described above then applies, unless the hypothesis of rational commensurability is satisfied. For details see [9].

The third construction, used to show that the threshold p_θ is always strictly positive and can be arbitrarily close to $3/2$ for rational θ, relies on Cantor sets. Let r, s be any two relatively prime positive integers, and consider the operator $T(f)(2x) = \int_{|t|<1} \prod_{j=1}^{3} f(\nu_j \cdot (x,t))\,dt$ where $\nu_1 + \nu_2 = (1,0)$ and $r^{-1}\nu_1 + s^{-1}\nu_2 = \nu_3$. Consider the following sets, which depend on r, s and on a large parameter K. Let

$$A = \{\sum_{n=1}^{K} a_n (rs)^{-n} + z\}$$

where each $a_n \in \{0, r, 2r, \dots, (s-1)r\}$ and z is an arbitrary real number satisfying $|z| \le 2(rs)^{-K}$, and likewise

$$B = \{\sum_{n=1}^{K} b_n (rs)^{-n} + z\}$$

where each $b_n \in \{0, s, 2s, \dots, (r-1)s\}$ and $|z| \le 2(rs)^{-K}\}$. Set $C = r^{-1}A + s^{-1}B$.

Then $T(\chi_A, \chi_B, \chi_C)(x) \gtrsim (rs)^{-K}$ whenever x is within distance $(rs)^{-K}$ of any sum $\sum_{n=1}^{K} (i_n r + j_n s)(rs)^{-n}$ with $0 \le i_n < s$ and $0 \le j_n < r$. Moreover, any two such sums are distinct, so the set of all of them has cardinality $(rs)^K$, and any two such sums differ by at least $(rs)^{-K}$. Letting $K \to \infty$, one finds that a necessary condition for T to map $(L^p)^3$ boundedly to L^q is that $1 \le r^{1-\frac{2}{p}} s^{1-\frac{2}{p}} (r+s)^{\frac{1}{p}}$. Details may be found in [9].

Arithmetic progressions again play a central role here; the Cantor sets in this construction should be regarded as projections onto $\mathbb{R}^1$ of (small δ-neighborhoods of) high-dimensional arithmetic multiprogressions, under mappings of the form $(x_0, \dots, x_K) \mapsto \sum_{j=0}^{K} (rs)^{-j} x_j$, for large K and small ε. This construction is related to Kahane's construction [26] of a Kakeya set, as well as to certain examples in [27].

3.7. **A positive result.**

3.7.1. *Setup for the proof.* The proof for the rationally commensurate case is an adaptation from the discrete to the continuum setting of an argument of Katz and Tao [27], who sharpened an estimate of Bourgain [3], descended from an argument of Gowers [19]. Their result directly rules out the existence of any examples resembling those constructed in the preceding section, but does not seem to directly imply any positive L^p inequalities. Instead, we will deduce from the method of proof in [27] continuum analogues of *variants* of its conclusion[13].

We continue to assume that $d = 1$. It suffices to show that there exist $\delta \in (0,1)$ and $0 < r_1, r_2, s$, depending on $\{\nu_j\}$ and satisfying $r_1 + r_2 + s = 2 - \delta$, such that for any

[13]The conclusion of [3, 27] concerns the cardinality of a certain image set; the variant we require also takes into account cardinalities of sets of preimages. Compare Theorem 3.2 with Corollary 3.3.

measurable sets $A, B, C \subset \mathbb{R}$ and any $\lambda > 0$, the set $E = \{x : T(\chi_A, \chi_B, \chi_C)(x) > \lambda\}$ satisfies

$$\lambda^{1-\delta}|E| \lesssim |A|^{r_1}|B|^{r_2}|C|^s. \tag{56}$$

Indeed, T maps $L^2 \otimes L^2 \otimes L^1$ boundedly to L^1, and so there is always the relatively trivial bound

$$\lambda|E| \lesssim \min\left(|A|^{1/2}|B|^{1/2}|C|^1, |A|^{1/2}|C|^{1/2}|B|^1, |B|^{1/2}|C|^{1/2}|A|^1\right). \tag{57}$$

Raising (56) to a power $\theta \in [0,1]$ and multiplying by any powers, summing to $1-\theta$, of the three inequalities (57), produces another inequality of the form (56). Choosing the powers appropriately yields an inequality (56) with $r_1 = r_2 = s$, albeit with a smaller δ. Writing it as a restricted weak type estimate $\lambda^{q_0}|E| \le (|A| \cdot |B| \cdot |C|)^{\frac{q_0}{p_0}}$, we find that $q_0 < 1$ and $\frac{3q_0}{p_0} = 1 + q_0$, whence $p_0 < 3/2$. From multilinear interpolation[14] [25, 28] we then conclude easily that for any $p > p_0$, T maps the threefold tensor product of L^p boundedly to L^q where $\frac{3}{p} = 1 + \frac{1}{q}$.

For the trilinear case in dimension one, the simplest operator of our family is $T_\infty(f,g,h)(x) = \int_{|t|<1} f(x+t)g(x-t)h(t)\,dt$. We will discuss in detail the analysis of this operator, proving that it maps $(L^p)^3$ to L^q, with $1 + q^{-1} = 3p^{-1}$, for all[15] $p > \frac{7}{5}$ (which is $< \frac{3}{2}$). Here the argument is based rather directly on [27].

By replacing B by $-B$, C by $\frac{1}{2}C = \{s/2 : s \in C\}$, and E by $\frac{1}{2}E$, we may arrange that for each $x \in E$,

$$|\{t \in C : t + x \in 2A \text{ and } t - x \in 2B\}| \ge \lambda. \tag{58}$$

Let

$$\tilde{G} = \{(a,b) \in A \times B : a + b \in C \text{ and } a - b \in E\}\ . \tag{59}$$

Select a measurable subset $G \subset \tilde{G}$ such that for each $x \in E$,

$$|\{(a,b) \in G : a - b = x\}| = \lambda,$$

and denote the latter set by G_x. Here the measure in question is one-dimensional Lebesgue measure on a line, determined by x, in $\mathbb{R}^2$. From this and the definition of $\tilde{G}$, it follows that (up to a constant numerical factor)

$$|G| = \lambda|E|. \tag{60}$$

3.7.2. *Lower bounds for inflations.* The reasoning will require lower bounds of the following type, which are most easily explained in terms of finite sets. Let E, Ω be finite sets and $\pi : E \mapsto \Omega$ be any mapping. Define $E^\dagger = \{(e,e') \in E \times E : \pi(e) = \pi(e')\}$. We will call $E^\dagger$ an *inflation* of E; one can think of E as being fibered over Ω and of $E^\dagger$ as being defined by expanding each fiber (thereby squaring its cardinality).

Denote the cardinality of a set S by $|S|$. Then

$$|E^\dagger| \ge |E|^2/|\Omega|. \tag{61}$$

Proof. For $\omega \in \Omega$ define $\rho(\omega) = |\{e \in E : \pi(e) = \omega\}|$. Then $|E| = \sum_{\omega\in\Omega} \rho(\omega)$, while $|E^\dagger| = \sum_\omega \rho^2(\omega)$. Cauchy-Schwarz thus gives $|E| \le |\Omega|^{1/2}|E^\dagger|^{1/2}$. □

[14] See also [21] for a nice account of multilinear interpolation with weak-type estimates for exponents less than one, based on Wolff's proof of the reiteration theorem.

[15] It seems improbable that this exponent 7/5 is optimal.

This result implies a variant of itself: The set of all $e \in E$ for which $\rho(\pi(e)) \geq \frac{1}{2}|E|/|\Omega|$ has cardinality $\geq |E|/2$. Indeed, let F be the set of all e which fail to satisfy this condition. Since $|F^\dagger| \leq \sum_{e\in F} \rho(\pi(e))$, we have $|F^\dagger| < |F| \cdot \frac{1}{2}|E|/|\Omega|$. On the other hand, by (61), $|F|^2/|\Omega| \leq |F^\dagger|$. Combining these conclusions gives $|F| < \frac{1}{2}|E|$. □

Thus whereas (61) merely yields an average lower bound of $|E|/|\Omega$ for the multiplicity function $e \mapsto \rho(\pi(e))$, we have found that this function cannot be concentrated too strongly on too small a subset of E.

3.7.3. *A merely heuristic argument.* Consider the extreme case, where $G = A\times B$, $A+B \subset C$, and for some $\lambda > 0$ one has $|\{(a,b) \in A\times B : a-b=e\}| \sim \lambda$, uniformly for all $e \in E$. Then $\lambda|E| \sim |G| = |A|\cdot|B|$.

Let S' be the set of all $((a_1,b_1),(a_2,b_2)) \in G\times G$ satisfying $a_1-b_1 = a_2-b_2$. Then $|S'| \lesssim \lambda^2|E|$; here $|S'|$ denotes the three-dimensional measure of S'. This bound holds because for any $e\in E$, the set of all $(a,b)\in G$ satisfying $a-b=e$ has (one-dimensional) measure $\lesssim \lambda$, whence the set of all $((a_1,b_1),(a_2,b_2))$ satisfying $a_1-b_1 = e = a_2-b_2$ has (two-dimensional) measure $\lesssim \lambda^2$.

On the other hand, consider $S = \{((a,b),(a',b')) \in G\times G : a+b = a'+b'\}$. By the continuum analogue of (61), $|S| \gtrsim |G|^2/|C|$ since $A+B\subset C$. Now whenever $((a,b),(a',b')) \in S$, we have $a-b' = a'-b$. The image of S under the mapping $((a,b),(a',b')) \mapsto ((a,b'),(a',b))$ therefore has (three-dimensional Hausdorff) measure $= |S| \gtrsim |G|^2/|C|$, but is contained in S' and therefore has measure $\lesssim \lambda^2|E|$.

Since $\lambda|E| \lesssim |G|$, we have
$$\lambda^2|E|^2 \lesssim |G|^2 \lesssim |C|\cdot|S| \lesssim |C|\lambda^2|E|,$$
whence $|E| \lesssim |C|$. We are given that $\lambda|E| \lesssim |A|\cdot|B|$; multiplying this with the preceding inequality gives $\lambda|E|^2 \lesssim |A|\cdot|B|\cdot|C|$, which is a restricted weak type inequality of type $(L^p)^3 \mapsto L^q$ with $p=1$ and $q=1/2$.

3.7.4. *Two applications of the inflation bound.* Returning to the proof of the theorem, introduce

$$V = \{(a,b,b') \in \mathbb{R}^3 : (a,b)\in G \text{ and } (a,b')\in G\}. \tag{62}$$

V is an inflation of G; G is fibered over A via the projection $(a,b)\mapsto a$, and V is obtained by replacing each fiber F by $F\times F$.

For $|V|$ there is a lower bound

$$|V| \geq |G|^2/|A|. \tag{63}$$

Here $|\cdot|$ denotes of course the Lebesgue measure of a set in the appropriate dimension, not cardinality. (63) is proved by the same reasoning used above for inflations of finite sets. Indeed, consider the map $\pi : G\mapsto A$ defined by $\pi(a,b) = a$. Set $\rho(a) = |\pi^{-1}\{a\}|$, for $a\in A$, where $|\cdot|$ denotes one-dimensional measure on the fibers. Then
$$|V| = \int_A \rho(a)^2\,da \geq |A|^{-1}\Big(\int_A \rho\Big)^2 = |A|^{-1}|G|^2.$$

Next, define $S \subset V^4$ to be the set of all points (v_0,v_1,v_2,v_3) satisfying all of the following relations

$$\begin{aligned} a_0+b_0 &= a_1+b_1, & a_0+b_0' &= a_1+b_1',\\ b_1 &= b_2, & b_1' &= b_2',\\ a_2+b_2 &= a_3+b_3, & b_2' &= b_3'. \end{aligned} \tag{64}$$

S is contained in a linear subspace of $\mathbb{R}^{12}$ of dimension 6. As coordinates for S, we use $V \times A \times B^2 \ni (v_0, a_2, b_3, b_3') = (a_0, b_0, b_0', a_2, b_3, b_3')$. All other coordinates a_j, b_j, b_j' of any point of S may be determined from these six coordinates by various linear relations.[16] Thus a simple upper bound for the size of S is $|S| \lesssim |A| \cdot |B|^2 \cdot |V|$.

The proof uses the basic technique of counting the same set in two different ways; we will derive both an upper bound and a lower bound for $|S|$, by different means, then compare the two results. For the set S there is the lower bound

$$|S| \gtrsim |V|^4/|B|^3|C|^3 \;, \tag{65}$$

which implies

$$|S| \gtrsim \frac{|G|^8}{|A|^4|B|^3|C|^3} \gtrsim \frac{\lambda^8|E|^8}{|A|^4|B|^3|C|^3}.$$

(65) may be proved by iterating the inflation bound, in its strong pointwise form. Write $V^4 = V_0 \times V_1 \times V_2 \times V_3$ where each V_j is a copy of V. Let S_1 be the set of all $(v_0, v_1) \in V_0 \times V_1$ satisfying the two relations on the first line of (64); then $|S_1| \gtrsim |V|^2/|C|^2$ by the fundamental lower bound on inflations. As in the finite case, not only does $|S_1|$ satisfy this lower bound, but the associated multiplicity function is $\gtrsim |V|/|C|^2$ on at least half of V, in the sense of Lebesgue measure. Thus we may replace V throughout the argument by such a favorable subset.

Now consider $V_1 \times V_2$, and the subset S_2 consisting of all points satisfying the relations on the second line of (64). Repeat the same reasoning, to obtain a subset of V having measure $\geq |V|/4$ on which the multiplicity functions associated to S_1, S_2 both are $\gtrsim$ their average values at every point. Once again, replace V by that subset. Repeating this reasoning one last time for $S_3 \subset V_2 \times V_3$ yields the desired bound. The details are left to the reader. Alternatively, one may consult Lemma 2.1 of [27].

3.7.5. *An improved upper bound for* $|S|$. The key point of the analysis is that a special algebraic relation holds, allowing a sharpening of the simple upper bound $|S| \lesssim |A| \cdot |B|^2 \cdot |V|$, as follows. If $(v_0, v_1, v_2, v_3) = s \in S$, then

$$a_3 - b_3' = a_2 - b_3 + b_0 - b_0'. \tag{66}$$

Indeed,

$$\begin{aligned}
a_3 - b_3' &= a_3 - b_2' + (b_2' - b_3') \\
&= (a_2 + b_2 - b_3) - b_2' + [(a_3 + b_3) - (a_2 + b_2)] + (b_2' - b_3') \\
&= (a_2 - b_3) + (b_2 - b_2') + [(a_3 + b_3) - (a_2 + b_2)] + (b_2' - b_3') \\
&= (a_2 - b_3) + (b_1 - b_1') + (b_2 - b_1) - (b_2' - b_1') + [(a_3 + b_3) - (a_2 + b_2)] + (b_2' - b_3') \\
&= a_2 - b_3 + b_0 - b_0' \\
&\qquad - [(a_0 + b_0) - (a_0 + b_0')] + (b_2 - b_1) - (b_2' - b_1') + [(a_3 + b_3) - (a_2 + b_2)] + (b_2' - b_3').
\end{aligned}$$

for any $(v_0, v_1, v_2, v_3) \in S$, the relations (64) imply that the second line of the final expression is a sum of terms vanishing identically on S, so the $a_3 - b_3'$ simplifies to $(a_2 - b_3) + (b_0 - b_0')$.

Thus once the coordinates a_2, b_0, b_0', b_3 of a point s of S are known, $a_3 - b_3'$ can be computed. But since $\varphi(s) = (a_3, b_3')$ belongs to G by construction, $\varphi(s) \in G_x$, where $x = a_2 + b_0 - b_0' - b_3$. Projecting $\varphi(s)$ onto its second coordinate, we conclude that for

[16] Knowing a_0, b_0, b_0' together with the relations defining S gives $a_1 + b_1, a_1 + b_1'$. This in turn gives $a_1 + b_2, a_1 + b_2'$. Since b_2' equals the coordinate b_3', this determines a_1, hence b_2 as well; b_1, b_1' are then determined, and finally a_3 is determined by the relation $a_3 = a_2 + b_2 - b_3$.

any (a_2, b_0, b_0', b_3), for any $s \in S$ sharing these four coordinates, the coordinate b_3' of s must belong to the projection of G_x onto the b_3' axis, and the measure of this projection is $\le \sqrt{2}\cdot\lambda$. Thus for any point of S, the coordinate b_3' is constrained to lie in a set determined by (a_2, b_0, b_0', b_3), and having measure $\lesssim \lambda$. Thus we arrive at the crucial estimate

$$|S| \lesssim \lambda |A| \cdot |B| \cdot |V| \ . \tag{67}$$

3.7.6. *Conclusion of the proof.* Combining (65) with (67), we deduce that

$$\lambda |A| \cdot |B| \cdot |V| \gtrsim |V|^4 |B|^{-3} |C|^{-3}, \tag{68}$$

so

$$|V| \lesssim \lambda^{1/3} |A|^{1/3} |B|^{4/3} |C|.$$

Since $|V| \gtrsim |G|^2/|A|$, this yields

$$|G| \lesssim \lambda^{1/6} |A|^{2/3} |B|^{2/3} |C|^{1/2}.$$

Since $|G| = \lambda |E|$, this can be rewritten as

$$\lambda^{5/6} |E| \lesssim |A|^{2/3} |B|^{2/3} |C|^{1/2}. \tag{69}$$

If the roles of A, B, C could be symmetrized, this would then yield $\lambda^{5/6}|E| \lesssim (|A| \cdot |B| \cdot |C|)^{11/18}$, which would yield exponents $p = 15/11$ and $q = 5/6$ in our final conclusion. But such a symmetrization does not seem to follow directly from the above setup, and since there is little reason to believe that $p = 15/11$ is optimal, we have not labored to obtain it. Instead, the above discussion of (56) yields a conclusion of the desired form, for all $p > 7/5$. This concludes the discussion of the simplest case in which estimates do hold with $p < 3/2$. $\square$

3.7.7. *The general rationally commensurate case.* For $d = 1$, any rationally commensurate trilinear operator in our class can be reduced to one of the operators $T_\theta(f, g, h)(t) = \int_{-1}^{1} f(x_t) g(x - t) h(x - \theta t)\, dt$ with $\theta \in \mathbb{Q}$, by linear changes of variables and rescalings of f, g, h. The proof [10] is based on the same ideas as for T_∞, but G is inflated to an auxiliary space $\tilde{V} \subset A \times B^L$ for some power L rather than to $V \subset A \times B^2$, and V^4 is replaced by $\tilde{V}^M$ for some power M, where L, M depend on the numerator and denominator of θ in an apparently complicated manner, and both can be arbitrarily large. Likewise, a large number of relations analogous to (64) are imposed on $\tilde{V}^M$ to define a subset S which has the decisive property that a suitable linear combination of two variables, (of which one is a coordinate for S) which takes values in E and may be expressed as a nontrivial linear combination of other coordinates of S. As in the special case treated above, this leads to an improved upper bound on $|S|$, with one factor of either $|A|, |B|$, or $|C|$ replaced by $C\lambda$.

Suppose that what is given is that for any $(a, b) \in G$, one has $a + b \in C$ and $ma - nb \in E$; and for instance that $m > n > 0$ and m, n are relatively prime. Then the main step is an algebraic construction which essentially reduces the pair of coefficients (m, n) to $(m - n, n)$. For details see [10]. There seems little point in presenting these details here, for I don't know how to do so in an illuminating way.

References

[1] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal. **1** (1991), 147–187.

[2] J. Bourgain, *Averages in the plane over convex curves and maximal operators*, J. Analyse Math. **47** (1986), 69–85.

[3] J. Bourgain, *On the dimension of Kakeya sets and related maximal inequalities*, Geom. Funct. Anal. 9 (1999), 256-282.
[4] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. 9 (1999), no. 5, 968–984.
[5] H. J. Brascamp, E. H. Lieb, and J. M. Luttinger, *A general rearrangement inequality for multiple integrals*, J. Functional Analysis 17 (1974), 227–237.
[6] A. Carbery, M. Christ and J. Wright, *Multidimensional van der Corput and sublevel set estimates*, J. Amer. Math. Soc. 12 (1999), no. 4, 981–1015.
[7] M. Christ, *Estimates for the k-plane transform*, Indiana Univ. Math. J. 33(1984), 891-910.
[8] M. Christ, *Hilbert transforms along curves, I: Nilpotent groups*, *Annals of Math.* **122** (1985), 575–596.
[9] M. Christ, *Convolution, Curvature, and Combinatorics. A Case Study*, International Math. Research Notices 19 (1998), 1033-1048.
[10] M. Christ, *On certain elementary trilinear operators*, Math. Research Letters 8 (2001), 43-56.
[11] M. Christ and M. Burak Erdoğan, *Mixed norm estimates for a restricted X-ray transform*, to appear, Journal D'Analyse.
[12] M. Christ and A. Kiselev, *Maximal functions associated to filtrations*, J. Funct. Anal. 179 (2001), 409-425.
[13] M. Christ and A. Kiselev, *WKB Asymptotic Behavior of Almost All Generalized Eigenfunctions for One-Dimensional Schrödinger Operators With Slowly Decaying Potentials*, J. Funct. Anal. 179 (2001), 426-447.
[14] M. Christ and A. Kiselev, Scattering and wave operators for one-dimensional Schrödinger operators with slowly decaying nonsmooth potentials, Geom. Funct. Anal., to appear.
[15] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, *Singular and maximal Radon transforms: Analysis and geometry*, Annals of Math. 150 (1999), 489-577.
[16] S. W. Drury, *L^p estimates for the x-ray transform*, Illinois J. Math. 27 (1983), 125-129.
[17] S. W. Drury, *Generalizations of Riesz potentials and L^p estimates for certain k-plane transforms*, Illinois J. Math. 28 (1984), no. 3, 495–512.
[18] M. B. Erdoğan, *Mixed norm estimates for a restricted X-ray transform in $\mathbb{R}^4$ and $\mathbb{R}^5$*, Internat. Math. Res. Notices 2001, no. 11, 575–600.
[19] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), no. 3, 529–551.
[20] L. Grafakos and N. Kalton, *Some remarks on multilinear maps and interpolation*, Math. Ann. 319 (2001), no. 1, 151–180.
[21] L. Grafakos and T. Tao, *Multilinear interpolation between adjoint operators*, preprint.
[22] A. Greenleaf and A. Seeger, *Fourier integral operators with cusp singularities*, Amer. J. Math. 120 (1998), no. 5, 1077–1119.
[23] A. Greenleaf, A. Seeger, and S. Wainger, *On X-ray transforms for rigid line complexes and integrals over curves in R^4*, Proc. Amer. Math. Soc. 127 (1999), no. 12, 3533–3545.
[24] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, fifth edition, Oxford University Press, Oxford, 1985.
[25] S. Janson, *On interpolation of multilinear operators*, Function spaces and applications (Lund, 1986), 290–302, Lecture Notes in Math., 1302, Springer, Berlin-New York, 1988.
[26] J. P. Kahane, *Trois notes sur les ensembles parfaits linéaires*, L'Enseignement Math. 15 (1969), 185-192.
[27] N. Katz and T. Tao, *Bounds on arithmetic projections, and applications to the Kakeya conjecture*, Math. Res. Letters 6 (1999), 625-630.
[28] C. Kenig and E. M Stein, *Multilinear estimates and fractional integration*, Math. Res. Letters 6 (1999), 1–15.
[29] M. Lacey and C. Thiele, *On Calderón's conjecture*, Ann. of Math. (2) 149 (1999), no. 2, 475–496.
[30] M. Lacey and C. Thiele, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$*, Ann. of Math. (2) 146 (1997), no. 3, 693–724.
[31] E. H. Lieb and M. Loss, *Analysis*, Amer. Math. Soc., Providence, 1997.
[32] Matoušek, Jiři *Lectures on discrete geometry*, Graduate Texts in Mathematics, 212. Springer-Verlag, New York, 2002.
[33] G. Mockenhaupt, S. Seeger, and C. D. Sogge, *Wave front sets, local smoothing and Bourgain's circular maximal theorem*, Ann. of Math. (2) 136 (1992), no. 1, 207–218.
[34] C. Muscalu, T. Tao, and C. Thiele, *Multilinear operators given by singular multipliers*, J. Amer. Math. Soc. 15 (2002), 469-496.

[35] D. Oberlin, Convolution estimates for some measures on curves, *Proc. Amer. Math. Soc.* **99** (1987), 56–60.
[36] D. Oberlin, A convolution estimate for a measure on a curve in $\mathbb{R}^4$, *Proc. Amer. Math. Soc.* **125** (1997), 1355–1361.
[37] D. Oberlin, *A convolution estimate for a measure on a curve in* $\mathbb{R}^4$, Proc. Amer. Math. Soc. 127 (1999), no. 1, 217–221.
[38] D. Oberlin, H. F. Smith, and C. D. Sogge, *Averages over curves with torsion*, Math. Res. Lett. 5 (1998), no. 4, 535–539.
[39] D.H. Phong and E. M. Stein, *Hilbert integrals, singular integrals and Radon transforms I*, Acta Math. **157** (1986), 99-157.
[40] I. Z. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arithmetica 60 (1991), 191-202.
[41] I. Z. Ruzsa, *Generalized arithmetical progressions and sumsets*, Acta Math. Hungar. 65 (1994), no. 4, 379–388.
[42] I. Z. Ruzsa, *Sums of finite sets*, Number theory (New York, 1991–1995), 281–293, Springer, New York, 1996.
[43] W. Schlag, *A geometric proof of the circular maximal theorem*, Duke Math. J. 93 (1998), no. 3, 505–533.
[44] W. Schlag, *A generalization of Bourgain's circular maximal theorem*, J. Amer. Math. Soc. **10** (1997), 103–122.
[45] W. Schlag, *On continuum incidence problems related to harmonic analysis*, preprint.
[46] W. Schlag and C. D. Sogge, *Local smoothing estimates related to the circular maximal theorem*, Math. Res. Lett. 4 (1997), no. 1, 1–15.
[47] A. Seeger, *Degenerate Fourier integral operators in the plane*, Duke Math. J. 71 (1993), no. 3, 685–745.
[48] A. Seeger, *Radon transforms and finite type conditions*, J. Amer. Math. Soc. 11 (1998), no. 4, 869–897.
[49] I. R. Shafarevich, *Basic Algebraic Geometry*, Springer-Verlag, Berlin, 1977.
[50] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.
[51] T. Tao and J. Wright, L^p *improving estimates for averages along curves*, preprint.
[52] T. Wolff, An improved bound for Kakeya type maximal functions, *Rev. Mat. Iberoamericana* **11** (1995), 651–674.
[53] T. Wolff, A Kakeya type problem for circles, *Amer. J. Math.* **119** (1997), 985–1026.
[54] T. Wolff, *Recent work connected with the Kakeya problem*, Prospects in mathematics (Princeton, NJ, 1996), 129–162, Amer. Math. Soc., Providence, RI, 1999.
[55] T. Wolff, *A mixed norm estimate for the X-ray transform*, Revista Mat. Iberoamericana 14 (1998), 561-601.
[56] T. Wolff, *A sharp bilinear cone restriction estimate*, Ann. of Math. (2) 153 (2001), no. 3, 661–698.
[57] T. Wolff, *A note on interpolation spaces*, Harmonic analysis (Minneapolis, Minn., 1981), pp. 199–204, Lecture Notes in Math., 908, Springer, Berlin-New York, 1982.

Michael Christ, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA

E-mail address: `mchrist@math.berkeley.edu`