A Sharpened Hausdorff-Young Inequality

Michael Christ
University of California, Berkeley

IPAM Workshop
Kakeya Problem, Restriction Problem, Sum-Product Theory
and perhaps more
May 5, 2014

Hausdorff-Young inequality

- Fourier transform normalized to be unitary on $L^2(\mathbb{R}^d)$:

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

- $\|\widehat{f}\|_{L^\infty} \le \|f\|_{L^1}$.
- Hausdorff-Young inequality $\boxed{\|\widehat{f}\|_{L^q} \le \|f\|_{L^p}}$ whenever $1 \le p \le 2$, $q = p' = \frac{p}{p-1}$ = conjugate exponent.
- Inequality valid on any locally compact Abelian group.
- Optimal constant $= 1$ for many groups, including $\mathbb{T}^d$, $\mathbb{Z}^d$.

Sharp constant and extremizers

- For $\mathbb{R}^d$, $\|\widehat{f}\|_{L^q} \leq \mathbb{A}_p^d \|f\|_{L^p}$ with optimal constant

$$\mathbb{A}_\mathbf{p} = \mathbf{p}^{1/2\mathbf{p}}\, \mathbf{q}^{-1/2\mathbf{q}} \; < \; \mathbf{1}.$$

- Babenko [1961] for $q = 4, 6, 8, 10, \ldots$;
  Beckner [1975] for all $p \in (1, 2)$.
- All **Gaussian** functions are extremizers;

$$\mathbf{G(x) = ce^{Q(x)+x \cdot v}}$$

  where Q is a negative definite homogeneous *real*–valued quadratic polynomial, and $v \in \mathbb{C}^d$.
- Lieb [1990] showed that there are no other extremizers.

Non-Quantitative Theorem

- Notation: $\mathfrak{G}$ = set of all Gaussians.
(a finite-dimensional subvariety of $L^p(\mathbb{R}^d)$)
- **Theorem Epsilon.** For every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|\widehat{f}\|_q \geq (1-\delta)\mathbb{A}_p^d \|f\|_p \implies \text{distance}(f, \mathfrak{G}) \leq \varepsilon \|f\|_p.$$

This is a **compactness** theorem; see below.
(Distance = distance in L^p.)

Symmetries

The inequality has a large noncompact group of symmetries; $\Phi(f) = \frac{\|\widehat{f}\|_q}{\|f\|_p}$ is unchanged under:

- $f \mapsto f \circ A$ **for** $A \in \mathrm{Gl}(d)$**.** This includes changes of scale $f(x) \mapsto f(rx)$.
- Translation
- Modulation $f \mapsto e^{iv\cdot x}$ for $v \in \mathbb{R}^d$
- Multiplication by scalars.

Compactness

- Theorem Epsilon is equivalent to a compactness statement:

  Let a sequence (f_n) satisfy

  $$\begin{cases} \|f_n\|_p \to 1 \\ \|\widehat{f_n}\|_q \to \mathbb{A}_p^d. \end{cases}$$

  Then some **renormalized** subsequence $(f^\star_{n_\nu})$ converges in L^p norm.

- *Renormalized* means that each $f_n^\star$ is obtained from f_n by action of an element of the symmetry group, and $\|f_n^\star\|_p \equiv 1$.

Main Theorem

- There exists $c > 0$ such that for every nonzero real-valued function $f \in L^p(\mathbb{R}^d)$,

$$\|\widehat{f}\|_q \leq \mathbb{A}_p^d \|f\|_p - \mathbf{c}\|\mathbf{f}\|_{\mathbf{p}}^{-1} \,\text{distance}(\mathbf{f}, \mathfrak{G})^{\mathbf{2}}.$$

Main Theorem

- There exists $c > 0$ such that for every nonzero real-valued function $f \in L^p(\mathbb{R}^d)$,

$$\|\widehat{f}\|_q \leq \mathbb{A}_p^d \|f\|_p - \mathbf{c}\|\mathbf{f}\|_{\mathbf{p}}^{-1} \, \text{distance}(\mathbf{f}, \mathfrak{G})^{\mathbf{2}}.$$

- If $\text{distance}(f, \mathfrak{G})/\|f\|_p$ is sufficiently small then

$$\|\widehat{f}\|_q \leq \mathbb{A}_p^d \|f\|_p - \mathbb{B}_{\mathbf{p},\mathbf{d}}\|\mathbf{f}\|_{\mathbf{p}}^{-1} \, \text{distance}(\mathbf{f}, \mathfrak{G})^{\mathbf{2}} + o\Big(\|f\|_p^{-1} \, \text{distance}(f, \mathfrak{G})\Big)^2 \|f\|_p$$

- where $\boxed{\mathbb{B}_{p,d} = \frac{1}{2}(p-1)(2-p)p^{d(2-p)/2p}\mathbb{A}_p^d}$

Optimal constant (in modified statement)

There is a more precise and stronger formulation, involving a different measurement of distance to $\mathfrak{G}$, in which the constant

$$\mathbb{B}_{p,d} = \frac{1}{2}(p-1)(2-p)p^{d(2-p)/2p}\mathbb{A}_p^d$$

in the theorem is optimal

unless I've made algebraic errors in the calculation. The analysis, in the hands of a competent calculator, certainly gives an explicit optimal constant.

Logic

- Theorem Epsilon is an essential step towards the Main Theorem.
- Once one knows that a near-extremizer is close to the set $\mathfrak{G}$ of Gaussians, one can completely switch tactics, employing Taylor expansion of the function $\|\widehat{f}\|_q / \|f\|_p$.
- Most of the effort goes towards the proof of Theorem Epsilon.

Multiprogressions

- Discrete multiprogression:
  $Q = \{u + \sum_{i=1}^{\mathbf{r}} n_i v_i : 0 \leq n_i i < N_i\}$.
  Rank of $Q \;=\; \mathbf{r}$.
  Q is proper if this mapping $\vec{n} \mapsto \ldots$ is injective.

- Continuum multiprogression: $P = Q + K$ in $\mathbb{R}^d$ where

$$\begin{cases} K \text{ is any ellipsoid} \\ Q \text{ is a discrete multiprogression.} \end{cases}$$

- Henceforth: "Progression" = "Multiprogression".

Quasi-extremizers

- For an inequality

$$\|Tf\|_q \leq A\|f\|_p,$$

a δ–quasi-extremizer is f satisfying

$$\|Tf\|_q \geq \delta \|f\|_p.$$

Quasi-extremizers

- For an inequality

$$\|Tf\|_q \le A\|f\|_p,$$

a δ–quasi-extremizer is f satisfying

$$\|Tf\|_q \ge \delta \|f\|_p.$$

- δ is permitted to be arbitrarily small, and will often be understood, so that one speaks simply of a quasi-extremizer.
- It is often of interest to characterize quasi-extremizers.
- Doing so, for arbitrarily *small* δ, is one of several key steps in the proof of Theorem Epsilon.

Connection with Young's Convolution Inequality

- Young's inequality: $\|f * g\|_r \leq B\|f\|_s\|g\|_s$
- **Lemma.** For any $\delta > 0$: If $\|\widehat{f}\|_q \geq \delta\|f\|_p$ then

$$\left\||f|^\gamma * |f|^\gamma\right\|_r \geq c\delta^2\|f\|_s^2$$

  where γ, r, s are determined by p.

- **Sharp** constants for the two inequalities are **not** connected in this way.

Young's inequality and Sumsets

We all know that these are closely related.

- Functions are weighted sets.
- Convolutions are weighted sumsets.

Quasi-extremizers for Young's inequality

Lemma. Suppose $\|f * f\|_r \geq \delta \|f\|_p^2$. Then there exists a disjointly supported decomposition

$$\boxed{f = g + h}$$ where h **is smaller than** f and g **has structure**:

- $\|h\|_p \leq (1 - c\delta^\gamma)\|f\|_p$
- there exists a (continuum) **multiprogression** P such that

$$g \text{ is supported on } P$$

$$|g(x)| \underset{\delta}{\asymp} |P|^{-1/p} \text{ where } g(x) \neq 0.$$

$$P \text{ has rank} \leq C_\delta$$

1D Brunn-Minkowski Inequality

- The Brunn-Minkowski inequalities encapsulate a fundamental property of algebraic sums of sets, in various contexts. They provide lower bounds for $|A+B|$ in terms of $|A|, |B|$.

- For (Borel measurable) $A, B \subset \mathbb{R}^1$,

$$|\mathbf{A}+\mathbf{B}| \geq |\mathbf{A}| + |\mathbf{B}|$$

where $|\cdot| =$ Lebesgue measure.

Equality only for intervals minus null sets.

Stable Converses to Brunn-Minkowski

- **Little Theorem.** Let $A \subset \mathbb{Z}$ be a finite set. If $|A+A| < 3|A| - 3$ then A is contained in an arithmetic progression P satisfying $|P| \leq |A+A| - |A| + 1$.

- **Big Theorem.** If A, B are finite sets of comparable cardinalities and if $|A+B| \leq K|A|$ then $A \subset P$ for some **multi**progression P of **rank** $\leq C_K$ and **cardinality** $\leq C_K|A|$.

- One can easily pass to the limit to obtain corresponding results for subsets of $\mathbb{R}$ (for the Little Theorem) and $\mathbb{R}^d$ (for the Big Theorem), with Lebesgue measure replacing cardinality.

Balog-Szemeredi Theorem

Let $A, B \subset \mathbb{Z}^d$ have comparable measures. If

$$\|\mathbf{1}_A * \mathbf{1}_B\|_{L^2}^2 \geq K^{-1} |A|^{3/2} |B|^{3/2}$$

then there exist subsets A', B' whose measures are significant fractions of $|A|, |B|$ such that

$$|A' + B'| \leq C_K |A'| + C_K |B'|.$$

Connection with Young's convolution inequality

Corollary: If $A, B \subset \mathbb{R}^d$ are sets with positive, finite Lebesgue measures, if their measures are comparable, and if

$$\|\mathbf{1}_A * \mathbf{1}_B\|_2^2 \geq \delta |A|^{3/2} |B|^{3/2},$$

then there exists a progression P such that

$$|P| \leq C_\delta |A|,$$
$$|P \cap A| \geq c_\delta |A|,$$
$$P \text{ has rank } \leq C_\delta.$$

Two facts about general operators/inequalities

Let $T : L^p \to L^q$ be a bounded linear operator with norm A. Suppose

$$\|T(f)\|_q \geq (1-\delta)A\|f\|_p$$
$$f = \varphi + \psi \text{ with disjoint supports.}$$

Then:

- $\|\varphi\|_p \geq \eta\|f\|_p \implies \|T(\varphi)\|_q \geq c\eta^C\|\varphi\|_p$
- $\min(\|\varphi\|_p, \|\psi\|_p) \geq \eta\|f\|_p \implies \|T(\varphi)T(\psi)\|_{q/2} \geq c\eta^C\|f\|_p^2$

  provided δ is sufficiently small relative to η.

Structure of Near-Extremizers — Preliminary

Proposition. Suppose $\boxed{\|\widehat{f}\|_q \geq (1-\delta)\mathbb{A}_p^d \|f\|_p}$. Let $\varepsilon > 0$. If δ is sufficiently small relative to ε then there exist a disjointly supported decomposition $f = g + h$ and $\boxed{\text{one}}$ **progression** P satisfying

$$\|h\|_p < \varepsilon \|f\|_p$$

and

$$\begin{cases} g \text{ is supported on } P \\ \|g\|_\infty |P|^{1/p} \leq C_\varepsilon \|f\|_p \\ P \text{ has rank} \leq C_\varepsilon. \end{cases}$$

A lemma used in preceding Proposition

Let exponents s, r be related so that $L^s * L^s \subset L^r$.

Lemma. Let P, Q be progressions of ranks $\leq \mathbf{r}$ satisfying and comparable Lebesgue measures. Suppose that

$$\|\mathbf{1}_P * \mathbf{1}_Q\|_r \geq \eta |P|^{1/s} |Q|^{1/s}.$$

Then there exists a progression R of rank $\leq C_{\mathbf{r}}$ such that

$$|R| \leq C|P|$$
$$P + Q \subset R.$$

End of the Beginning

- I have a series of papers in which additive combinatorics — more specifically, Freiman's Little Theorem — is used to prove compactness. Inequalities previously treated in this way are
  1. Young's convolution inequality,
  2. the Brunn-Minkowski inequality,
  3. the Riesz-Sobolev inequality.

- But I have not found a way to apply the Little Theorem to the Hausdorff-Young inequality.

- Likewise, symmetrization/rearrangement inequalities do not seem to be applicable.

Time to get down to brass tacks

- It was clear early on that Freiman's Big Theorem could be used to obtain a weaker version of the Proposition stated two slides above. But how can one show that the discrete part of the progression P has rank 0?

- Reasoning so far has been heavily, though not exclusively, of form

$$\textit{Nearly zero} \Longrightarrow \textit{not optimally large.}$$

- When near optimality has been exploited, it has not used specific properties of Hausdorff-Young / Fourier transform.

- Need to better exploit near optimality.

Key Step Remaining

Multiprogression P can be replaced by a convex set K.

(see subsequent slides)

A stroke of good fortune

$\widehat{f}|\widehat{f}|^{q-2}$ is also a near-extremizer of Hausdorff-Young and therefore everything proved so far about f applies to $\widehat{f}$ (in L^q norm);

$\widehat{f}$ is nearly concentrated on a convex set L.

Other Remaining Steps

- $|K| \cdot |L| \leq C_\varepsilon$ (provided δ is sufficiently small).
- Likewise, K, L are contained in dual ellipsoids up to translation and multiplication by a uniformly bounded factor.
- Symmetries of the equation can be invoked to make K, L be balls centered at 0 with uniformly bounded radii.
- A hitch: $\|\widehat{h}\|_q$ small does not imply any useful bound for h. A tiny additional argument based on strict convexity of the unit ball of L^p gives compactness.

The case of discrete groups

- Optimal constant in Hausdorff-Young for $\mathbb{Z}^m$ equals 1.
- f is extremizer if and only if f supported on a single point.
- **Theorem.** If $\|\widehat{f}\|_q \geq (1-\delta)\|f\|_p$ then there exists $z \in \mathbb{Z}^m$ such that
$$\|f\|_{L^p(\mathbb{Z}^m \setminus \{\mathbf{z}\})} < \varepsilon \|f\|_p,$$
where $\varepsilon = \varepsilon(\delta, p) \to 0$ as $\delta \to 0$.
- Theorem due to Eisner–Tao [2012], and subsequently Charalambides–Christ [2011 preprint; non-Abelian case]. A key idea already in Fournier [1977].
- Proof of [C-C] rests on $|A + B| \geq |A| + |B| - 1$.

Main Idea for completion of Theorem Epsilon

- Sparse progressions are the enemy.
- **Will prove:**
  If f lives almost entirely on a very sparse progression P then the situation is nearly that of a discrete group.
- Contradiction! (in light of preceding Theorem).

Any progression can be approximated by $\mathbb{Z}^d$

- Consider $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$.
- Define $\|x\|_{\mathbb{R}^d/\mathbb{Z}^d} =$ distance from x to $\mathbb{Z}^d$.
- **Lemma.** Let P be a progression in $\mathbb{R}^d$ of any finite rank $\mathbf{r}$ satisfying $|P| = 1$. Let $\delta \le \frac{1}{2}$.
  There exists an affine automorphism $\mathcal{T}$ of $\mathbb{R}^d$ such that

  $$\|\mathcal{T}(\mathbf{x})\|_{\mathbb{R}^\mathbf{d}/\mathbb{Z}^\mathbf{d}} < \delta \quad \text{for all } x \in P$$

  satisfying nondegeneracy condition

  $$|\text{Jacobian determinant}| \ge c\delta^{d\mathbf{r}+d^2}.$$

- This applies to P of arbitrarily high rank.

A clarification

During the talk, a distinguished member of the audience asked whether there should be an upper bound on the Jacobian determinant.

No upper bound is intended; the lemma is not true if any upper bound, even one that depends on δ, is imposed.

In the argument below, when this lemma is applied to near-extremizers, a contradiction is eventually obtained in cases in which the Jacobian is large.

A detour

- The diameter of a set in $\mathbb{R}^1$ is at least as large as its Lebesgue measure.
  The lemma on approximation of sparse progressions by $\mathbb{Z}^d$ under affine transformations required a higher-dimensional analogue.
- Let $\mathcal{M}_d$ be $d \times d$ real matrices.

  Lemma. Let $E \subset \mathcal{M}_d$ have positive, finite Lebesgue measure. There exist $A, B \in E$ such that

  $$|\det(A - B)| \geq c_d |E|^{1/d}.$$

- *I've actually proved only a slightly weaker statement with* $\det(\sum_{i=1}^{N(d)}(A_i - B_i))$*, which suffices for the application.*

Four Fourier transforms

- $\mathbb{R}^d$ Fourier transform $\widehat{\cdot}$
- $\mathbb{Z}^d$–Fourier transform $\mathfrak{F}$

$$\mathfrak{F}(f)(\theta) = \sum_{n \in \mathbb{Z}^d} f(n) e^{-2\pi i n \cdot \theta}$$

- $\mathbb{Z}^d \times \mathbb{R}^d$ partial Fourier transform $\mathcal{F}$

$$\mathcal{F}(F)(\theta, x) = \sum_{n \in \mathbb{Z}^d} F(n, x) e^{-2\pi i n \cdot \theta}.$$

  This is simply $\mathfrak{F}(F_x)$, where $F_x(n) = F(n, x)$.
- $\mathbb{Z}^d \times \mathbb{R}^d$ Fourier transform $\mathfrak{F}^{\times}$.

Lifting to $\mathbb{Z}^d \times \mathbb{R}^d$

- Lift f to F on $\mathbb{Z}^d \times \mathbb{R}^d$:

$$F(n, x) = f(n + x) \text{ if } |x - n| \leq \eta, \text{ and 0 otherwise.}$$

Lifting to $\mathbb{Z}^d \times \mathbb{R}^d$

- Lift f to F on $\mathbb{Z}^d \times \mathbb{R}^d$:

$$F(n,x) = f(n+x) \text{ if } |x-n| \le \eta, \text{ and 0 otherwise.}$$

- **Lemma.** If

$$\begin{cases} \|\widehat{f}\|_q \ge (1-\delta)\mathbb{A}_p^d \|f\|_p \\ f \text{ is supported within distance } \eta \text{ of } \mathbb{Z}^d \end{cases}$$

then $\quad \|\mathfrak{F}^\times(F)\|_{L^q_{\xi,\theta}} \ge (1-\delta - \mathbf{C}\eta^\gamma)\mathbb{A}_p^d \|F\|_{L^p_{x,n}}$

and therefore

$$\|\mathcal{F}(F)\|_{\mathbf{L^p_x L^q_\theta}} \ge (1-\delta - \mathbf{C}\eta^\gamma)\|F\|_{L^p_x L^p_n}.$$

(Recall: $\mathcal{F}$ is partial Fourier transform in the discrete variable.)

Explanation

Proof of the Lemma is based on a property of the lifting: For $\theta \in \mathbb{T}^d, \; m \in \mathbb{Z}^d, \; \alpha \in [-\eta, \eta]^d$,

$$\mathfrak{F}^{\times}(F)(\theta, m+\alpha) \underset{\eta \ll 1}{\approx} \widehat{f}(m+\theta).$$

(Fourier transform of a function with small diameter support is slowly varying.)

Explanation

Proof of the Lemma is based on a property of the lifting: For $\theta \in \mathbb{T}^d,\ m \in \mathbb{Z}^d,\ \alpha \in [-\eta, \eta]^d$,

$$\mathfrak{F}^{\times}(F)(\theta, m + \alpha) \underset{\eta \ll 1}{\approx} \widehat{f}(m + \theta).$$

(Fourier transform of a function with small diameter support is slowly varying.)

plus two properties exploited by Beckner and then Lieb:

- $\mathfrak{F}^{\times}$ is a composition of two partial Fourier transforms, with respect to n and x.
- Whenever $s \geq t$, $\quad \|F\|_{L^s_\xi L^t_n} \leq \|F\|_{L^t_n L^s_\xi}$.

Beginning of the End

- This concludes the outline of the proof of Theorem Epsilon.
- Begin fine analysis of functions close in norm to set $\mathfrak{G}$ of Gaussians.
- Use Taylor expansion about a Gaussian of the functional

$$\boldsymbol{\Phi}(\mathbf{f}) = \frac{\|\widehat{\mathbf{f}}\|_{\mathbf{q}}}{\|\mathbf{f}\|_{\mathbf{p}}}.$$

- Numerator is C^2 only in a limited sense. Denominator is simply not twice differentiable.
- The difficulty: A small *norm* perturbation is not small *pointwise*, and moreover, is assuredly not small pointwise relative to e^{-cx^2}.

Taylor expansion of functional

- Let $1 < \mathbf{p} < \mathbf{2} \le q$ and $T : L^p \to L^q$ bounded linear operator.
- Let G be a nonzero element of the manifold $\mathfrak{G}$ of extremizers.
- For any small $\varepsilon > 0$, for any f with sufficiently small norm orthogonal to the tangent space to $\mathfrak{G}$ at G (in appropriate sense), there is a disjointly supported splitting

$$f = f_\sharp + f_\flat$$

- such that $\Phi(h) = \|Th\|_q / \|h\|_p$ and its formal second variation $\mathcal{Q}$ satisfy

$$\Phi(\mathbf{G} + \mathbf{f}) \le \|T\| + \mathcal{Q}(f_\sharp) - \mathbf{c}_\varepsilon(\|f_\flat\|_p)^{\mathbf{p}} + \varepsilon \|f\|_p^2.$$

- Term $-c_\varepsilon \|f_\flat\|_p^p$ is enormous but sign is favorable.

Spectrum of an operator / Analysis of $\mathcal{Q}$

- Consider dimension $d = 1$ for simplicity.
- Define $G(x) = e^{-\pi x^2}$ and $G^t(x) = e^{-t\pi x^2}$.
- Let $s, t > 0$. Define compact positive self-adjoint operator $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$:

$$\mathbf{Tf} = \mathbf{G^t} \cdot \big(\mathbf{G^s} * (\mathbf{G^t f})\big).$$

- Analysis requires that the 4–th largest eigenvalue of T be less than a specific threshold, where s, t are determined by p. (Because Gaussians form a manifold of dimension 3.)
- **Eigenvalues of T can be computed in closed form** using Mehler kernel/Ornstein-Uhlenbeck semigroup.