Title
Multilinear and Sharpened Inequalities

Author
O'Neill, Kevin

Publication Date
2019

Peer reviewed|Thesis/dissertation

Multilinear and Sharpened Inequalities

by

Kevin William O'Neill

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael Christ, Chair
Professor Daniel Tataru
Professor Emeritus Kenneth Wachter

Spring 2019

Multilinear and Sharpened Inequalities

Abstract

Multilinear and Sharpened Inequalities

by

Kevin William O'Neill

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Michael Christ, Chair

Multilinear integral inequalities, such as Hölder's inequality and Young's convolution inequality, play a large role in analysis. In [6] and [5], Bennett, Carbery, Christ, and Tao provide a classification of such inequalities of the form

$$\int_{\mathbb{R}^d} \prod_{j=1}^{n} f_j(L_j(x))dx \leq C \prod_{j=1}^{n} ||f_j||_{p_j}$$

for constant $C > 0$, exponents $p_j \in [1, \infty]$, and surjective linear maps $L_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$, with $f_j : \mathbb{R}^{d_j} \to \mathbb{R}$.

In Chapter 2, we discuss a generalization of the above related to work by Ivanisvili and Volberg [23]. Specifically, we provide a classification of functions $B : \mathbb{R}^n \to [0, \infty)$ such that

$$\int_{\mathbb{R}^d} B(f_1(L_1(x)), ..., f_n(L_n(x))dx \leq CB\left(\int f_1, ..., \int f_n\right).$$

In some cases, it will be shown that maximizers of the above inequality exist. Tuples of Gaussians are not always maximizers, differing from the usual multilinear theory.

Chapter 3 focuses its attention on the trilinear form for twisted convolution:

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x)g(y)h(x+y)e^{i\sigma(x,y)}dxdy.$$

While existence of maximizers can shed light on the structure of an operator, sometimes it is useful to establish more refined information. For twisted convolution, we show a quantitative version of the statement that if a triple of functions nearly maximizes the form, then it must be close to a maximizing triple. Such a statement

may be referred to as a *sharpened inequality.* Here, the proof of a sharpened inequality is complicated by the fact that no maximizers exist for twisted convolution; however, one may vary the amount of oscillation and compare to the case in which there is zero oscillation.

In Chapter 4, we establish a sharpened version of the following inequality due to Baernstein and Taylor [3]:

$$\iint_{S^d\times S^d} f(x)g(y)h(x\cdot y)d\sigma(x)d\sigma(y) \leq \iint_{S^d\times S^d} f^*(x)g^*(y)h(x\cdot y)d\sigma(x)d\sigma(y)$$

where f, g, h are restricted to the class of indicator functions and h is monotonic on $[-1, 1]$. In the above, f^* refers to the symmetric decreasing rearrangement of f, and likewise for g and g^*.

To everyone who has taught me about math and about life.

Contents

Acknowledgments

First and foremost, I would like to thank my advisor, Michael Christ, for all his support during this project: for his mathematical guidance, patience, clear explanations, assistance with the job market, and understanding of work-life balance when it was most important.

I could not have done this without the support of my parents, who from the very beginning gave me a space where I could learn without judgment and focus on the intrinsic rewards over the extrinsic.

At Berkeley, I would like to thank Professors Daniel Tataru and Kenneth Wachter for serving on my dissertation committee. Also, thank you to Alexis Drouot, Justin Brereton, and Albert Ai for co-organizing the Harmonic Analysis and Differential Equations Seminar with me.

Outside academia, I would like to thank Andrew for supporting me from the beginning, Brad for surviving a six-year roommate relationship with me, Nyralee for listening to me talk about my research more than anyone who wasn't my advisor, the 2013 Stags for helping me develop my character, and to all the above for keeping me sane during this process.

Chapter 1

Introduction to Extremizers and Multilinear Forms

Given a linear operator $T : L^{p_1}(\mathbb{R}^d) \to L^{p_2}(\mathbb{R}^d)$, it is often useful to prove estimates of the form

$$||Tf||_{p_2} \leq C||f||_{p_1} \tag{1.1}$$

for some $0 < C < \infty$. For one, this implies T is continuous. However, it is sometimes desirable to obtain further information about T.

Question 1.1. What is the optimal constant C such that (1.1) holds?

For linear operators such as the above, this optimal constant is the operator norm, though we will often ask the question for inequalities of a more complicated form.

Question 1.2. If A is said optimal constant, does there exist nonzero $g \in L^{p_1}$ such that $||Tg||_{p_2} = A||g||_{p_1}$? If so, can we characterize such g?

When such a g exists, we refer to it as a *maximizer* (or *extremizer*) of T. If T is a bounded linear operator, then there always exists an optimal constant C in (1.1) by the least upper bound property of the real numbers. However, not all such operators have maximizers.

For example, let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ by $(a_n) \mapsto ((1 - 2^{-n})a_n)$. Clearly, T has norm 1, yet there is no nonzero sequence $(b_n) \in \ell^2(\mathbb{N})$ such that $\sum_n |b_n|^2 = \sum_n (1 - 2^{-n})^2 |b_n|^2$. By compactness of the unit ball, maximizers trivially exist in the case of linear transformations on finite-dimensional domains. In the infinite-dimensional case, this reasoning may fail, yet it motivates the effort towards establishing precompactness, such as in Chapter 2.

One instance of a nontrivial operator which has maximizers is the Fourier transform $\mathcal{F} : L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ defined by

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) := \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} f(x)dx \tag{1.2}$$

whenever $p \in (1, 2)$. It was shown by Babenko [1] and Beckner [4] that

$$||\hat{f}||_{p'} \leq \mathbf{A}_p^d ||f||_p, \tag{1.3}$$

where $\mathbf{A}_p^d = \left[p^{1/2p}/p'^{1/2p'}\right]^d$ is the optimal constant, and later by Lieb [26] that the maximizers are precisely Gaussians.

We now use this example as an opportunity to discuss the role of symmetry in this type of analysis. Suppose that $g(x)$ is a maximizer of $\mathcal{F}$ and let $a > 0$ denote a constant. Then by the elementary identity $\widehat{e^{iax}g} = \hat{g}(\xi - a)$, we see that $e^{iax}g(x)$ is a maximizer as well, since L^p norms are invariant under translation and modulation.

For this reason, we refer to modulation– and any operation which preserves the ratio $||\hat{f}||_{p'}/||f||_p$– as a *symmetry* of $\mathcal{F}$. The set of maximizers of $\mathcal{F}$ is closed under the symmetries of $\mathcal{F}$ by definition. This same property holds for any operator with its maximizers and symmetries, and provides a natural manifold structure to the set of maximizers.

There are a number of other questions one could ask about maximizers– if they are unique up to symmetry, if the operator satisfies certain precompactness properties– though we turn our attention to just one.

To motivate this question, consider the following. Let $\mathcal{M}$ be the manifold of maximizers for a bounded linear operator T. By continuity, if f is close to $\mathcal{M}$, then $||Tf||/||f||$ is close to $||Tg||/||g||$ for $g \in \mathcal{M}$. However, the converse of this statement fails in general.

We now consider one particular quantitative formulation of this converse.

Question 1.3. Given a linear operator $T : L^{p_1}(\mathbb{R}^d) \to L^{p_2}(\mathbb{R}^d)$ with manifold of maximizers $\mathcal{M}$ and optimal constant A, does there exist $c > 0$ such that

$$||Tf||_{p_2} \leq A||f||_{p_1} - c\text{dist}(f, \mathcal{M})^2 \tag{1.4}$$

for some natural, appropriately normalized definition of distance function dist?

When such a c exists, we refer to (1.4) as a *sharpened inequality*.

The particular form of (1.4) is reminiscent of a truncated Taylor series centered at the maximum of a function on $\mathbb{R}^d$. Furthermore, in examples, the exponent 2 is often shown to be sharp.

In some scenarios, it is desirable to consider the distance between the orbit of f under the symmetries of the operator to a single, fixed maximizer. In either case, it is notable that the distance function respects the symmetries of the operator since they do not change the left hand side of (1.4).

Each of the three above questions may be formulated for multilinear operators ($L^{p_1} \times \cdots \times L^{p_n} \to L^q$) and multilinear forms ($L^{p_1} \times \cdots \times L^{p_n} \to \mathbb{R}$). In this thesis, we will answer Question 1.3 for certain multilinear forms in Chapters 3 and 4. But before doing so, we digress to provide some context on multilinear forms.

1.1 Multilinear Forms and Inequalities

The field of analysis is abundant with multilinear integral inequalities. Long-known examples include:

- Multilinear Hölder's inequality

$$\left| \int_{\mathbb{R}^d} \prod_{j=1}^{n} f_j(x) dx \right| \leq \prod_{j=1}^{n} ||f_j||_{p_j}, \tag{1.5}$$

  whenever $\sum_{j=1}^{n} p_j^{-1} = 1$ and $p_j \in [1, \infty]$.

- Young's convolution inequality (in dual form)

$$\left| \int_{\mathbb{R}^d \times \mathbb{R}^d} f_1(x) f_2(y) f_3(x+y) dx dy \right| \leq \prod_{j=1}^{3} ||f_j||_{p_j}, \tag{1.6}$$

  for $\sum_{j=1}^{3} p_j^{-1} = 2$ and $p_j \in [1, \infty]$.

- Loomis-Whitney inequality

$$\left| \int_{\mathbb{R}^d} \prod_{j=1}^{d} f_j(x_1, ..., x_{j-1}, x_{j+1}, ..., x_d) dx \right| \leq \prod_{j=1}^{d} ||f_j||_{d-1}. \tag{1.7}$$

  Note the above only holds with all exponents equal to $d-1$.

While there are other examples of multilinear integral inequalities (such as the Hardy-Littlewood-Sobolev inequality and boundedness of the bilinear Hilbert transform), the three examples above are all of the form

$$\left|\int_{\mathbb{R}^d}\prod_j f_j\circ L_j(x)dx\right|\leq C\prod_j ||f_j||_{p_j} \tag{1.8}$$

for surjective maps $L_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ and some $C > 0$. In particular, the range of allowable exponents is the set of (p_j) satisfying certain linear inequalities.

The work of Bennett-Carbery-Christ-Tao classified all instances of such inequalities (see [5,6]). Specifically, there exists a $C > 0$ such that (1.8) holds if and only if both

$$\sum_{j=1}^{n}\frac{d_j}{p_j}=d \tag{1.9}$$

and

$$\dim(V)\leq\sum_{j=1}^{n}\frac{\dim(L_jV)}{p_j} \tag{1.10}$$

for all vector subspaces $V \subset \mathbb{R}^d$. These inequalities are known as the *Hölder-Brascamp-Lieb inequalities.*

In [5], the authors completely answer Question 1.1 and provide a nearly complete answer for Question 1.2 for Hölder-Brascamp-Lieb inequalities. Note that here, one is concerned with maximizing n-tuples of functions $(g_1, ..., g_n)$.

In Chapter 2, we introduce a generalization of the Hölder-Brascamp-Lieb inequalities and classify all such examples. We then answer Question 1.2 for a particular subclass of examples. While we are able to show maximizers exist, it remains open what those maximizers are. However, it is shown they are not always the usual suspects (tuples of Gaussians).

The work of Christ [17] answers Question 1.3 in the case of Young's convolution inequality (1.6), and in Chapter 3 we extend this work to the trilinear operator for twisted convolution

$$\mathcal{T}'(\mathbf{f}) := \iint_{\mathbb{R}^{2d}\times\mathbb{R}^{2d}} f_1(x)f_2(y)f_3(x+y)e^{i\sigma(x,y)}dxdy, \tag{1.11}$$

where σ is the symplectic form on $\mathbb{R}^{2d}$.

1.2 Rearrangement Inequalities and Restriction to Indicator Functions

Another way of illuminating which functions have larger outputs under a particular operator or form may sometimes be found in a rearrangement inequality. For this purpose, given a function $f : \mathbb{R}^d \to [0, \infty)$, we define its *nonincreasing symmetric rearrangement* as the (unique) upper semi-continuous, radially symmetric, decreasing function $f^* : \mathbb{R}^d \to [0, \infty)$ such that $|\{x : f(x) > \lambda\}| = |\{x : f^*(x) > \lambda\}|$ for all $\lambda > 0$. (Here, by radial, we mean $f(x) = g(|x|)$ for some $g : [0, \infty) \to [0, \infty)$ and by nonincreasing, we mean g is nonincreasing.)

The classical Riesz-Sobolev inequality [9, 28, 30] states that for arbitrary $f, g, h : \mathbb{R}^d \to [0, \infty)$,

$$\iint f(x)g(y)h(x+y)dxdy \leq \iint f^*(x)g^*(y)h^*(x+y)dxdy. \tag{1.12}$$

See [9, 10] for discussions of the history of the inequality.

While this is an interesting result in its own right, it has found applications in the analysis of maximizers of integral inequalities. For instance, [12] uses the Riesz-Sobolev inequality in its proof of qualitative stability for Young's convolution inequality, and [25] uses a related result in establishing the existence of maximizers for the Hardy-Littlewood-Sobolev inequality. In Chapter 2, we use a generalization of (1.12) to prove the existence of maximizers in a generalization of Young's convolution inequality.

One feature of the Riesz-Sobolev inequality is that it lends naturally to the restriction to indicator functions. For instance, proofs of the inequality (such as that featured in [27]) work via "layer-cake decomposition," reducing to proving (1.12) for indicator functions. Also, letting $\mathbf{1}_E$ denote the indicator function of the set $E \subset \mathbb{R}^d$, $(\mathbf{1}_E)^* = \mathbf{1}_{E^*}$, where E^* is the closed ball of measure $|E|$ centered at the origin. Lastly, indicator functions are more natural for our purpose of proving sharpened inequalities.

To this point, we have only considered function spaces on $\mathbb{R}^d$. However, some of this theory may be generalized to other domains, including the Riesz-Sobolev inequality. Letting f^* denote the symmetric, decreasing arrangement of $f : S^d \to [0, \infty)$ (see Chapter 4 for a precise definition for functions on the sphere), one has the following generalization of the Riesz-Sobolev inequality, due to Baernstein-Taylor [3].

Let $f, g : S^d \to [0, \infty)$ and let $h : [-1, 1] \to \mathbb{R}$ be a monotonically increasing function. Then,

$$\iint_{S^d \times S^d} f(x)g(y)h(x \cdot y)dxdy \leq \iint_{S^d \times S^d} f^*(x)g^*(y)h(x \cdot y)d\sigma(x)d\sigma(y), \qquad (1.13)$$

where σ denotes the surface measure on S^d.

In the case $d = 1$, one may check that the integrals in (1.13) correspond to the trilinear form for convolution defined in terms of the natural group structure on S^1. In the case of Abelian groups, Christ and Iliopoulou proved a sharpened inequality for (1.13) [18]. The higher-dimensional case requires a different sort of analysis and is established in Chapter 4.

Chapter 2

A Variation on Hölder-Brascamp-Lieb Inequalities

2.1 Introduction

In a dual form, Young's convolution inequality on $\mathbb{R}^d$ states that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y)g(x-y)h(x)dxdy \leq C_{p,q,r,d}||f||_p||g||_q||h||_r, \tag{2.1}$$

where $p,q,r \in [1,\infty]$, $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=2$ (interpreting $1/\infty$ as 0) and $C_{p,q,r,d}$ is the optimal constant.

It was established in [4], [26], and [9] that certain compatible triples of Gaussians are the maximizers of (2.1), providing a sharp form of the inequality. Later, Carlen, Lieb, and Loss [11] proved this by running an L^p-norm-preserving heat equation through time with f, g, and h as initial data and showing that the left hand side is nondecreasing with time.

Recall from Chapter 1 that [6] provides the following generalization of Young's inequality which also encompasses Hölder's inequality and the Loomis-Whitney inequality. Let d, n, d_j be positive integers ($1 \leq j \leq n$) and let $L_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ be surjective linear maps. Then there exists $C < \infty$ such that

$$\int_{\mathbb{R}^d}\prod_{j=1}^{n} f_j(L_j(x))dx \leq C\prod_{j=1}^{n}||f_j||_{L^{p_j}(\mathbb{R}^{d_j})} \tag{2.2}$$

for all $f_j \in L^{p_j}(\mathbb{R}^{d_j})$ and with C depending only on d, n, d_j, and L_j if and only if both

$$\sum_{j=1}^{n} \frac{d_j}{p_j} = d \tag{2.3}$$

and

$$\dim(V) \leq \sum_{j=1}^{n} \frac{\dim(L_j V)}{p_j} \tag{2.4}$$

for all vector subspaces $V \subset \mathbb{R}^d$. The set of exponents $(1/p_1, ..., 1/p_n)$ satisfying both (2.3) and (2.4) is called the *Hölder-Brascamp-Lieb (HBL) polytope*. Thus, the HBL polytope is compact and convex with finitely many extreme points.

One may obtain (2.1) from (2.2) by setting $d = 2k, n = 3, d_j = k$, and $L_1(x, y) = y, L_2(x, y) = x - y, L_3(x, y) = x$, where $\mathbb{R}^{2k} = \{(x, y) : x, y \in \mathbb{R}^k\}$. In another paper ([6]), Bennett, Carbery, Christ, and Tao proved the existence of maximizers (in particular, certain tuples of Gaussians) by a generalization of the above heat equation method.

(2.2) may be rewritten in the form

$$\int_{\mathbb{R}^d} \prod_{j=1}^{n} f_j(L_j(x))^{s_j} dx \leq C \prod_{j=1}^{n} \left(\int_{\mathbb{R}^{d_j}} f_j \right)^{s_j}, \tag{2.5}$$

where $s_j = 1/p_j$ and $f_j \geq 0$. (This is a non-restricting assumption since $|\int f| \leq \int |f|$.) In this chapter, we will frequently use the notation $s = (s_1, ..., s_n)$. The above may be rewritten as

$$\int_{\mathbb{R}^d} B(f_1(L_1(x)), ..., f_n(L_n(x)))dx \leq CB\left(\int f_1, ..., \int f_n \right), \tag{2.6}$$

where $B(y_1, ..., y_n) = y_1^{s_1} \cdots y_n^{s_n}$. In this chapter, we will say $B : \mathbb{R}_+^n \to \mathbb{R}_+$ is a *Hölder-Brascamp-Lieb (HBL) function for* $\{L_j\}$ if (2.6) holds for all nonnegative $f_j \in L^1(\mathbb{R}^{d_j})$. Here $\mathbb{R}_+ = [0, \infty)$.

A similar question was explored in [23] in the case where each L_j is of rank 1 ($d_j \equiv 1$). The authors found sufficient conditions on B for the left hand side of (2.6) to be bounded by the same expression where the f_j are replaced with certain Gaussians G_j satisfying $\int f_j = \int G_j$. A corollary of this result is that certain tuples of Gaussians are among the extremizers. The key condition was a concavity requirement on B which allowed the heat equation method from [11] to work. Their bounding term matches ours in the case where each L_j is the identity.

In this chapter, we remove the rank 1 restriction and provide necessary and sufficient conditions for a function $B : \mathbb{R}_+^n \to \mathbb{R}_+$ to be an HBL function in the following theorem, which will be proven in Section 2.2. Part of the proof will involve the construction of a parallelipiped with particular dimensions through a dual linear programming problem as in [21].

By $A \lesssim B$, we mean that there exists a $0 < C < \infty$ such that $A \le CB$ and by $A \gtrsim B$, we mean there exists a $0 < C' < \infty$ such that $A \ge CB$. $A \approx B$ means $A \lesssim B$ and $A \gtrsim B$.

Theorem 2.1. *Let $B : \mathbb{R}_+^n \to \mathbb{R}_+$ be nondecreasing in each coordinate and satisfy $B(y_1, ..., y_n) = 0$ whenever any of the y_j are 0. Let $d, d_j, (1 \le j \le n)$ be positive integers and $L_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ surjective linear maps whose Hölder-Brascamp-Lieb polytope $\mathcal{P}$ is nonempty. Then the following are equivalent:*

(a) B is an HBL function for $\{L_j\}$.

(b) For all $0 < \lambda_j, y_j < \infty$,

$$B(\lambda_1 y_1, ..., \lambda_n y_n) \lesssim \max_{s \in \mathcal{P}} \lambda^{s_1} \cdots \lambda_n^{s_n} B(y_1, ..., y_n). \tag{2.7}$$

(c) For all $0 < \lambda_j, y_j < \infty$,

$$B(\lambda_1 y_1, ..., \lambda_n y_n) \gtrsim \min_{s \in \mathcal{P}} \lambda^{s_1} \cdots \lambda_n^{s_n} B(y_1, ..., y_n). \tag{2.8}$$

Allowing for a change of underlying constant, each of the possible conclusions in the above theorem is invariant under multiplication of B by a bounded function with bounded inverse. Thus, the theorem still holds if we replace the hypothesis that B is nondecreasing in each coordinate with the weaker hypothesis that B is bounded above and below by a positive multiple of a function which is nondecreasing in each coordinate.

Example 2.1. Given a collection of linear maps $\{L_j\}_{j=1}^n$ and $(s_1, ..., s_n)$ lying in its Hölder-Brascamp-Lieb polytope, $B(y_1, ..., y_n) = y_1^{s_1} \cdots y_n^{s_n}$ is an HBL function.

Example 2.2. The class of HBL functions for any $\{L_j\}$ is closed under linear combinations with nonnegative coefficients and taking minima and maxima. For instance, if $B_1, ..., B_4$ are HBL functions, then so is $\min\{\max\{B_1, B_2\}, 2B_3 + B_4\}$.

Here, the advantage of Theorem 2.1 is control over the implicit constants. For instance if B_1 and B_2 satisfy (2.7) with constant 1, then so does $\max\{B_1, B_2\}$.

The class of HBL functions is also closed under pointwise limits of sequences, provided there is uniformity in the implicit constants. In particular, one may combine different HBL functions with an integral. As an example, $B(y_1, y_2, y_3) =$

$\int_{-1/6}^{1/6} y_1^{2/3-t/2} y_2^{2/3-t/2} y_3^{2/3+t} dt$ is an HBL function for the linear maps found in Young's convolution inequality.

Example 2.3 (Uniqueness of Loomis-Whitney inequalities)**.** The classical Loomis-Whitney inequality is given by

$$\int_{\mathbb{R}^d} B(f_1(\pi_1(x)), ..., f_d(\pi_d(x)))dx \leq B\left(\int f_1, ... \int f_d\right), \tag{2.9}$$

where $B(y_1, ..., y_d) = y_1^{1/(d-1)} \cdots y_n^{1/(d-1)}$. Previous theory says this B is the only power function which also serves as an HBL function. Theorem 2.1 says, up to multiplication by constants, this B is the only HBL function (even allowing for non-power functions).

The remainder of the chapter is dedicated to the question of maximizers. In particular, we will focus on the choice of d, n, d_j, L_j used in Young's inequality to emphasize the differences in setting rather than prove statements in their most general form.

In Section 2.3, we will state and prove a rearrangement inequality that allows one to replace each f_j with its symmetric decreasing rearrangement. The proof of this uses the classical technique found in [20], where it was shown that $\int F(f(x), g(x))dx \leq \int F(f^*(x), g^*(x))dx$ for F satisfying a certain second-order condition.

In Section 2.4, we will show that for certain B, near-maximizing triples of (2.6) must be localized in scale and that these scales must be close for each function in the triple. This result is similar to the one found in [12] for the setting of L^p norms and will be used in establishing precompactness. Section 2.5 will piece together these arguments to establish the existence of extremizers in certain cases of HBL functions, as stated in the following theorem.

For notation, let $\vec{y} = (y_1, ..., y_n)$ denote a vector in $\mathbb{R}_+^n$ and let $\Delta_3(B; a, b, c, d, e, f)$ denote the third order difference:

$$\begin{aligned} &B(b, d, f) - B(a, d, f) - B(b, c, f) - B(b, d, e) \\ &\qquad + B(b, c, e) + B(a, d, e) + B(a, c, f) - B(a, c, e). \end{aligned} \tag{2.10}$$

Theorem 2.2. *Let $P_i(a, b, c) = a^{1/p_i} b^{1/q_i} c^{1/r_i}$, where $p_i, q_i, r_i \in (1, \infty)$ and $1/p_i + 1/q_i + 1/r_i = 2$. Let $B = \rho(P_1, ..., P_n)$ where it is assumed that*

$$\rho(\lambda_1 y_1, ..., \lambda_n y_n) \leq C \max_i \lambda_i \rho(y_1, ..., y_n)$$

for all $0 < \lambda_i, y_i < \infty$ and

$$\rho(\vec{y_1}) + \rho(\vec{y_2}) \leq \rho(\vec{y_1} + \vec{y_2})$$

for all $\vec{y_i} \in \mathbb{R}^n_+$. Furthermore, suppose B is continuous with

$$B(0,0,0) = B(x,0,0) = B(0,y,0) = B(0,0,z) = 0,$$

along with

$$\Delta_3(B; a, b, c, d, e, f) \geq 0$$

for all $a \leq b, c \leq d, e \leq f$.

Let $\alpha, \beta, \gamma > 0$. Then, there exist f, g, h which maximize

$$\iint B(f(y), g(x-y), h(x))dxdy$$

under the constraint $\int f = \alpha, \int g = \beta, \int h = \gamma$.

The setup of Theorem 2.2 includes the hypotheses of the rearrangement inequality from Section 2.3 as well as conditions which allow us to use some tools from the L^p norms setting while also extending the conclusion to other HBL functions.

Lastly, Section 2.6 will provide an example of an HBL function with non-Gaussian maximizers. We will prove this to be the case by showing that no Gaussian is a critical point with regards to the Euler-Lagrange equations and referencing the existence of extremizers result from Section 2.5.

2.2 Necessary and Sufficient Conditions for HBL functions

The proofs of $(c) \Rightarrow (b) \Rightarrow (a)$ are relatively straightforward so we will address those here before moving on to the proof of $(a) \Rightarrow (c)$.

Proof of $(c) \Rightarrow (b) \Rightarrow (a)$. Suppose (c) holds. Simultaneously replace each y_j in the given inequality with $\lambda_j y_j$ and each λ_j with λ_j^{-1}. Then (b) is obtained by dividing both sides by

$$\min_{s \in \mathcal{P}} \lambda^{-s_1} \cdots \lambda_n^{-s_n}$$

and then using the fact that the reciprocal of the minimum is the maximum of the reciprocals.

Now suppose (b) and consider nonnegative L^1 functions f_j. If any of the f_j has zero integral (hence is zero a.e.), then (2.6) holds trivially, so assume $\int f_j > 0$ for all

j. Letting $g_j(x) = \frac{f_j}{\int f_j}$, we rewrite the left hand side of the desired integral inequality to obtain

$$\int_{\mathbb{R}^d} B(f_1 \circ L_1(x), ..., f_n \circ L_n(x))dx$$
$$= \int_{\mathbb{R}^d} B\left(g_1 \circ L_1(x) \cdot \int f_1, ..., g_n \circ L_n(x) \cdot \int f_n\right) dx. \quad (2.11)$$

By applying (2.7), we may bound (2.11) by a constant times

$$\int_{\mathbb{R}^d} \max_{s \in \mathcal{P}} (g_1 \circ L_1(x))^{s_1} \cdots (g_n \circ L_n(x))^{s_n}\, dx \cdot B\left(\int f_1, ..., \int f_n\right).$$

Let us recall the fact that $\mathcal{P}$ is a compact, convex polytope. If $s, s' \in \mathcal{P}$, then taking any point on the segment between s and s' corresponds to taking a weighted geometric mean of $\lambda_1^{s_1} \cdots \lambda_n^{s_n}$ and $\lambda_1^{s_1'} \cdots \lambda_n^{s_n'}$. Thus, for any $x \in \mathbb{R}^d$, the above maximum may be obtained at extreme points of $\mathcal{P}$. We denote the set of extreme points of $\mathcal{P}$ as $\mathcal{P}'$. Since all terms are nonnegative, we may bound the maximum by a summation over extreme points to obtain

$$\int_{\mathbb{R}^d} \max_{s \in \mathcal{P}} (g_1 \circ L_1(x))^{s_1} \cdots (g_n \circ L_n(x))^{s_n}\, dx$$
$$\leq \int_{\mathbb{R}^d} \sum_{s \in \mathcal{P}'} (g_1 \circ L_1(x))^{s_1} \cdots (g_n \circ L_n(x))^{s_n}\, dx. \quad (2.12)$$

Next, we exchange the integral with the sum and bound each of the integral terms. Since each function g_n has integral equal to 1, we have

$$\int_{\mathbb{R}^d} (g_1 \circ L_1(x))^{s_1} \cdots (g_n \circ L_n(x))^{s_n}\, dx \leq C_s, \quad (2.13)$$

where C_s is the optimal constant such that

$$\int_{\mathbb{R}^d} \prod_{j=1}^{n} f_j(L_j(x))dx \leq C_s \prod_{j=1}^{n} ||f_j||_{L^{p_j}(\mathbb{R}^{d_j})}.$$

Since $\mathcal{P}$ has only finitely many extreme points, we combine (2.11), (2.12), and (2.13) to obtain

$$\int_{\mathbb{R}^k} B(f_1 \circ L_1(x), ..., f_n \circ L_n(x))dx \leq \left(\sum_{s \in \mathcal{P}'} C_s \right) B\left(\int f_1, ..., \int f_n \right)$$
$$= CB\left(\int f_1, ..., \int f_n \right).$$

□

The main goal of the remainder of the section will be to prove the following lemma.

Lemma 2.3. *Let* $\lambda = (\lambda_1, ..., \lambda_n)$ *such that* $\log \lambda_j$ *are nonnegative integers and let* $L_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ *be a collection of surjective linear maps with HBL polytope* $\mathcal{P}$ *as before. Then, there exists a parallelipiped* $S \subset \mathbb{R}^d$ *such that*

$$|S| \approx \min_{(s_1,...,s_n) \in \mathcal{P}} \lambda_1^{s_1} \cdots \lambda_n^{s_n}$$

and

$$|L_j(S)| \leq \lambda_j,$$

where the proportionality constants are independent of λ*.*

To see the usefulness of Lemma 2.3, let us demonstrate how it may be used to complete the proof of Theorem 2.1. The reduction to the case in which $\log \lambda_j$ are nonnegative integers will be established in Lemma 2.6.

Proof of $(a) \Rightarrow (c)$. Given λ_j such that $\log \lambda_j$ are nonnegative integers, let S be as in Lemma 2.3. Define $f_j = y_j 1_{L_j(S)}$. By plugging these f_j into (2.6), we obtain a left hand side equal to

$$|\cap_j L_j^{-1}(L_j(S))| B(y_1, ..., y_n) \geq |S| B(y_1, ..., y_n) = \min_{(s_1,...,s_n) \in \mathcal{P}} \lambda^{s_1} \cdots \lambda_n^{s_n} B(y_1, ..., y_n)$$

and a right hand side equal to

$$B(|L_1(S)| y_1, ..., |L_n(S)| y_n) \leq B(\lambda_1 y_1, ..., \lambda_n y_n).$$

Combining the two inequalities gives (2.8).

□

Now we begin the proof of Lemma 2.3. By taking logs of the minimum seen in (2.8), we reduce computing this term to a linear programming problem. Fixing $\lambda = (\lambda_1, ..., \lambda_n) \in \mathbb{R}^n_+$, we now define the *primal LPP* as

$$\text{minimize } \log \lambda \cdot s = \sum_j s_j \log \lambda_j \text{ over } s \in \mathbb{R}^n_+$$

subject to

$$\sum_j s_j \cdot \dim(L_j(V)) \geq \dim(V) \text{ for all } V \in \mathbf{E}, \qquad d = \sum_j s_j d_j, \qquad s_j \geq 0.$$

In the above, $\mathbf{E}$ is a finite list of subspaces which are sufficient to determine the HBL polytope. By this, we mean that (2.4) for only subspaces in $\mathbf{E}$ together with (2.3) is sufficient to describe $\mathcal{P}$. Because of this fact, we may add a finite number of subspaces to $\mathbf{E}$ without changing the optimum value of $\log \lambda \cdot s$.

One may note that while we have included the restriction $s_j \geq 0$, we have neglected to explicitly include the restriction $s_j \leq 1$. However, this may be obtained from the existing inequalities and proper choice of subspace as follows. Subtract the restriction $\dim V \leq \sum_j s_j \dim L_j(V)$ from $d = \sum_j s_j d_j$ to obtain

$$(d - \dim V) \geq \sum_j s_j (d_j - \dim L_j(V))$$

for all subspaces $V \subset \mathbb{R}^d$. Fix $1 \leq j_0 \leq n$ and pick $V = Ker(L_{j_0})$. By the Rank-Nullity theorem, the coefficient on s_{j_0} in the above is equal to $d - \dim V$. Since all other s_j are already taken to be nonnegative, $s_{j_0} \leq 1$. By taking $\mathbf{E}$ to include all subspaces of the form $Ker(L_j)$, we may recover the bounds $s_j \leq 1$.

Next, we prove three technical lemmas to aid us in the analysis of this linear programming problem. The first is preliminary, the second allows us to deal with only nonnegative solutions and coefficients, and the third will aid us in showing that a certain algorithm terminates.

Lemma 2.4. *If $B : \mathbb{R}^n_+ \to \mathbb{R}_+$ is an HBL function, then*

$$B(R^{d_1} y_1, ..., R^{d_n} y_n) \approx R^d B(y_1, ..., y_n)$$

for all $0 < R, y_j < \infty$.

Proof. Let $0 < R, y_j < \infty$ be arbitrary. Plug in the functions $f_j = y_j 1_{B_R(\mathbb{R}^{d_j})}$ to (2.6). The right hand side becomes $B(R^{d_1} y_1, ..., R^{d_n} y_n)$ while the left hand side scales like R^d, giving us the inequality

$$R^d B(y_1, ..., y_n) \lesssim B(R^{d_1} y_1, ..., R^{d_n} y_n).$$

Since the above holds for all $0 < R, y_j < \infty$, we may simultaneously replace R with $1/R$ and y_j with $R^{d_j} y_j$ to obtain the reverse inequality. □

Lemma 2.5. *It suffices to establish* (2.8) *for* $\lambda_j \geq 1$. *That is, if* $B : \mathbb{R}_+^n \to \mathbb{R}_+$ *is an HBL function and* (2.8) *holds for* $\lambda_j \geq 1$ *and* $0 < y_j < \infty$, *then it also holds for* $0 < \lambda_j, y_j < \infty$.

Proof. Let $0 < y_j, \lambda_j < \infty$ be given. Choose $R > 0$ sufficiently large such that $R^{d_j} \lambda_j > 1$ for all j. Then, by Lemma 2.4 and the fact that $d = \sum_j s_j d_j$ for any $s \in \mathcal{P}$,

$$\begin{aligned} R^d B(\lambda_1 y_1, ..., \lambda_n y_n) &\approx B(R^{d_1} \lambda_1 y_1, ..., R^{d_n} \lambda_n y_n) \\ &\gtrsim \min_{s \in \mathcal{P}} (R^{d_1} \lambda^{s_1}) \cdots (R^{d_n} \lambda_n^{s_n}) B(y_1, ..., y_n) \\ &= R^d \min_{s \in \mathcal{P}} \lambda_1^{s_1} \cdots \lambda_n^{s_n} B(y_1, ..., y_n). \end{aligned}$$

Dividing both sides by R^d gives the desired result. □

Lemma 2.6. *It suffices to establish* (2.8) *for* $\log \lambda_j \in \mathbb{N} \cup \{0\}$ *for all* j.

Proof. Choose nonnegative integers m_j such that $e^{m_j} \leq \lambda_j < e^{m_j + d_j}$. (We may take the $m_j \geq 0$ by the previous lemma.) Since B is nondecreasing in each coordinate, we have

$$B(e^{m_1} y_1, ..., e^{m_n} y_n) \leq B(\lambda_1 y_1, ..., \lambda_n y_n) \leq B(e^{m_1 + d_1} y_1, ..., e^{m_n + d_n} y_n).$$

By Lemma 2.4, all three terms are uniformly comparable up to a constant multiple of e^d. Similarly, for any $s \in \mathcal{P}$ (in particular the minimum),

$$\Pi_j (e^{m_j})^{s_j} \leq \Pi_j \lambda_j^{s_j} < \Pi_j (e^{m_j + d_j})^{s_j}.$$

Again, all three terms are equivalent up to a constant multiple of e^d by the relation $d = \sum_j s_j d_j$ for all $s \in \mathcal{P}$. By hypothesis, we have

$$B(e^{m_1} y_1, ..., e^{m_n} y_n) \gtrsim \min_{s \in \mathcal{P}} \Pi_j (e^{m_j})^{s_j} B(y_1, ..., y_n).$$

By replacing the above terms with the corresponding ones involving λ_j and adjusting the constant of proportionality, (2.8) for $\log \lambda_j \in \mathbb{N} \cup \{0\}$ extends to all $\lambda_j > 1$, and therefore all λ_j. □

Let $\dim(\mathbf{E}) = (\dim V)_{V\in\mathbf{E}}$. We define the *dual LPP* as

$$\text{maximize } y \cdot \dim(\mathbf{E})$$

subject to

$$y \cdot \dim(L_j(\mathbf{E})) \le \log \lambda_j \text{ for all } j, \qquad y_V \ge 0 \text{ for all } V \ne \mathbb{R}^d, \qquad y_{\mathbb{R}^d} \text{ free.}$$

The dual LPP relates to the primal LPP via the following elementary theorem from linear programming. For a source, see an introductory textbook on linear programming, such as [22].

Theorem 2.7 (Duality Theorem (special case))**.** *Let A be an $m\times n$ matrix, $c, x \in \mathbb{R}^n$, and $b, y \in \mathbb{R}^m$ for $m, n \ge 1$. Suppose that A, b, c have all nonnegative entries and $\{x : Ax \le b, x \ge 0\}$ is nonempty and bounded. Then, the maximum value of $c^T x$ subject to the constraints $Ax \le b, x \ge 0$ is equal to the minimum value of $y^T b$ subject to the constraints $y^T A \ge c^T, y \ge 0$. Furthermore, there exist optimal vectors x, y for both problems.*

By the above theorem, the optimal value of the dual LPP is equal to the optimal value of the primal LPP. In the remainder of this section, we will work with dual vectors y to construct a parellelipiped S whose volume is $e^{y\cdot\dim(\mathbf{E})}$. By taking the optimal value of $y \cdot \dim(\mathbf{E})$, we will show the volume of S is $\min_{s\in\mathcal{P}} \lambda^{s_1} \cdots \lambda_n^{s_n}$. We may then use S to construct functions f_j which we plug into (2.6) to obtain (2.8).

Since the remainder of this section will only involve the dual LPP with minimal reference to the primal LPP, we now make the following convention. Each dual vector y is of the form $(y_V)_{V\in\mathbf{V}}$, where $\mathbf{V}$ is the set of all subspaces of $\mathbb{R}^d$. If $\mathbf{W}$ is a collection of subspaces of $\mathbb{R}^d$, then we say a dual vector y is *supported on* $\boldsymbol{W}$ if $y_V = 0$ for all $V \notin \mathbf{W}$. Each vector y that we consider will be supported on a finite list of subspaces; hence the expression $y \cdot \mathbf{V}$ will always be well-defined.

To begin, we will show that y may be taken to be supported on a *flag*, which we define to be a sequence of properly nested subspaces $W_1 \subsetneq W_2 \subsetneq ... \subsetneq W_t = \mathbb{R}^d$.

Proposition 2.8. *Let y be an optimal dual vector of the dual LPP which is supported on $\mathbf{E}$. Then, there exists a dual vector y' supported on a flag such that $y \cdot \dim \mathbf{E} = y' \cdot \dim \mathbf{V}$ and $y' \cdot \dim(L_j(\mathbf{V})) \le y \cdot \dim(L_j(\mathbf{E})) \le \log \lambda_j$. Furthermore, there exists a finite list of subspaces $\mathbf{E}'$ independent of y such that y' may be chosen to be supported on $\mathbf{E}'$ for any optimal dual vector y.*

Before proving the Proposition 2.8, we remark that the finiteness of $\mathbf{E}'$ is advantageous for the following reason. When we construct the parallelipiped S, we would like the volumes of S and $L_j(S)$ to be porportional to the λ_j in appropriate ways. However, the proportionality constants will depend on the arrangement of the subspaces. A priori, if one changes λ_j, then one also changes the optimal dual vector, which changes which flag y' is supported on. But, limiting the subspaces to a finite list ensures that a single constant will work as the λ_j vary. This is nontrivial, since the algorithm developed in [21] involves summing and intersecting subspaces. It is known [8] that a finite list of subspaces will not necessarily generate a finite list under those operations. We work around this difficulty by performing these operations in a particular order and applying the following lemma.

Lemma 2.9. *Suppose $V \subset \mathbb{R}^d$ is a subspace and $W_1 \subset ... \subset W_t$ is a flag. Then $\{V, W_1, ..., W_t\}$ generates only a finite list of subspaces under the operations of repeated summation and intersection.*

Proof. It suffices to list all such subspaces and show the list is closed under summation and intersection. We claim the complete list is $\{V\} \cup \{W_i, V + W_i, V \cap W_i\}_{i=1}^t \cup \{W_i + (V \cap W_j)\}_{i<j}$.

Beginning with $\{V\} \cup \{W_i, V + W_i, V \cap W_i\}_{i=1}^t$, we note that most summations and intersections are already on this list since many subspaces are contained within one another and when $S \subset T$, we have $S + T = T$ and $S \cap T = S$. The two cases which this does not cover are $W_i + (V \cap W_j)$ where $i < j$ and $(V + W_i) \cap W_j$ where $i < j$. Since $W_i \subset W_j$, these two are equal and the last type of subspace on our list.

It remains to show that intersections and summations involving subspaces of the $W_i + (V \cap W_j)$ are still on our list. Adding two such subspaces, we find that

$$[W_{i_1} + (V \cap W_{j_1})] + [W_{i_2} + (V \cap W_{j_2})] = W_{\max(i_1,i_2)} + (V \cap W_{\max(j_1,j_2)}),$$

which is of the same form.

Similarly, intersecting two such subspaces, we find that

$$[(W_{i_1} + V) \cap W_{j_1}] \cap [(W_{i_2} + V) \cap W_{j_2}] = (W_{\min(i_1,i_2)} + V) \cap W_{\min(j_1,j_2)},$$

which is also of the same form. □

To prove the proposition, we will use the following *basic algorithm (BA)*: Given a vector y which is not supported on a flag, find two subspaces V and W such that neither is contained in the other and $y_V \geq y_W > 0$. Set $y'_{V+W} = y_{V+W} + y_W, y'_{V\cap W} = y_{V\cap W} + y_W, y'_W = 0, y'_V = y_V - y_W$. Repeat this process until the desired result.

It was shown in [21] that the BA terminates provided the initial y has all nonnegative and rational coordinates. Furthermore, at each step $y \cdot \dim(\mathbf{V})$ is preserved and $y \cdot \dim(L_j(\mathbf{V}))$ does not increase.

Proof of Proposition 2.8. Write $\mathbf{E} = (E_1, ..., E_k, \mathbb{R}^d)$. Perform the BA on y but only with respect to the coordinates y_{E_1} and y_{E_2}. This creates a flag $W_{1,1} \subsetneq ... \subsetneq W_{1,t_1}$ such that our modified y is supported on $\{W_{1,1}, ..., W_{1,t_1}, E_3, ..., E_k, \mathbb{R}^d\}$.

Now, given a y supported on a flag $W_{i,1} \subsetneq ... \subsetneq W_{i,t_i}$ and the remaining original subspaces $\{E_{i+2}, ..., E_k\}$, we perform the BA on y using only the subspaces $\{W_{i,1}, ..., W_{i,t_i}, E_{i+2}\}$. This converts y to a new dual vector supported on a flag $W_{i+1,1} \subsetneq ... \subsetneq W_{i+1,t_{i+1}}$ together with $E_{i+3}, ..., E_k$.

Continue this process until the list of subspaces E_i is exhausted, resulting in a dual vector supported solely on a flag. While $y_{\mathbb{R}^d}$ is excluded from modification, this does not prevent our final list from being a flag since every subspace is contained in $\mathbb{R}^d$.

Since the $\log \lambda_j$ are integers, we may take optimal y with all rational coordinates. In addition, each coordinate used in the BA is nonnegative as $y_{\mathbb{R}^d}$ is excluded from such operations. Since this algorithm is solely the concatenation of the BA performed on particular collections of subspaces and the BA is known to terminate in such an instance, our algorithm terminates.

It remains to prove the claim that a finite number of subspaces is considered. Certainly, in the case of a particular given y this is true as only finitely many subspaces are introduced in each of a finite number of steps. However, at each inductive step there are only finitely many subspaces which can be generated from the previous subspaces by Lemma 2.9. The total number of inductive steps is bounded by $k-1$, so the total number of subspaces may be counted via a finite tree.

□

Now we will begin the construction of particular functions which when plugged into (2.6) will establish (2.8).

Definition 2.4. Suppose a dual vector y is supported on an independent collection of subspaces $Y_1, ..., Y_t$ whose direct sum is $\mathbb{R}^d$. Define the parellipiped

$$S_y = \left\{ x \in \mathbb{R}^d | x = \sum_{i=1}^{t} \sum_{j=1}^{j_i} a_i^j v_i^j, 0 \le a_i^j \le e^{y_{Y_i}} \right\},$$

where $\{v_i^1, ..., v_i^{j_i}\}$ is a (fixed) basis for Y_i.

We cite the following two results from [21]. While they were proven in the context of Hölder-Brascamp-Lieb inequalities over the integers, the proofs for the results as stated here may be obtained by simply repeating the proofs from [21], but replacing $\mathbb{Z}$ with $\mathbb{R}$ and $\mathbb{Z}^d$ with $\mathbb{R}^d$. Similarly, the dependence on the subspaces Y_i may be deduced by simply following the proofs.

Let $\{V_i\}_{i=1}^n$ be a collection of subspaces of a vector space V. We say $\{V_i\}_{i=1}^n$ is linearly independent if $\{v_1, ..., v_n\}$ is linearly independent for all nonzero $v_i \in V_i$.

Proposition 2.10. *Let y be a dual vector supported on linearly independent subspaces $Y_1, ..., Y_t$ whose direct sum is $\mathbb{R}^d$. Then,*

$$|S_y| \approx e^{y\cdot \dim(\mathbf{V})},$$

where the proportionality constant depends only on the Y_i.

Lemma 2.11. *Let y be a dual vector supported on linearly independent subspaces $Y_1, ..., Y_t$ whose direct sum is $\mathbb{R}^d$. Let $W_i = Y_1 + ... + Y_i$.*

Let $L : \mathbb{R}^d \to \mathbb{R}^{d'}$ be any linear map and set $c_i = \dim(L(W_i)) - \dim(L(W_{i-1}))$. Then

$$|L(S_y)| \lesssim e^{\sum y_{Y_i} c_i},$$

where the proportionality constant depends only on L and the Y_i (or equivalently, the W_i).

Now fix y as the dual vector supported on a flag $W_1 \subsetneq ... \subsetneq W_t$ as obtained from Proposition 2.8. Choose linearly independent subspaces Y_i of W_i such that $Y_1 + ... + Y_i = W_i$ and define the dual vector y' supported on $\{Y_1, ..., Y_t\}$ by

$$y'_{Y_i} = y_{W_i} + ... + y_{W_t}. \tag{2.14}$$

Proof of Lemma 2.3. Fix a list of subspaces $\mathbf{E}$ which are sufficient to determine the HBL polytope and include $Ker(L_j)$ and all the subspaces generated in Proposition 2.8.

Let y be an optimal dual vector from the dual LPP, modified by Proposition 2.8 to be supported on a flag. Define $S = S_{y'}$, where y' is the dual vector obtained in (2.14). Then, by Proposition 2.10,

$$\begin{aligned} |S| \approx e^{y'\cdot \dim(\mathbf{E})} &= e^{\sum_i (y_{W_i} + ... + y_{W_t})(\dim Y_i)} \\ &= e^{\sum_i y_{W_i} (\dim Y_1 + ... + \dim Y_i)} = e^{\sum_i y_{W_i} \dim W_i}. \end{aligned}$$

Since y is an optimal dual vector, the value of $\sum_i y_{W_i} \dim W_i$ above is optimal and hence equal to the optimal value of $s \cdot \log \lambda$ from the primal LPP, giving us the desired volume estimate.

Similarly, by Lemma 2.11,

$$|L_i(S)| \approx e^{\sum_i y'_{Y_i} c_i} = e^{\sum_i (y_{W_i} + \ldots + y_{W_t}) c_i}$$
$$= e^{\sum_i y_{W_i}(c_1 + \ldots + c_i)} = e^{\sum_i y_{W_i} dim(L_j(W_i))} \leq e^{\log \lambda_j} = \lambda_j.$$

where the last step follows from the constraints on dual vectors. We may obtain $|L_j(S)| \leq \lambda_j$ in place of $|L_j(S)| \lesssim \lambda_j$ by a uniform scaling of S with scaling parameter dependent only on the previous proportionality constants.

□

2.3 Rearrangement Inequality

Given a function $f : \mathbb{R}^d \to \mathbb{R}$, let $E_f(\lambda) = |\{x \in \mathbb{R} : f(x) \geq \lambda\}|$ denote its distribution function. If $E_f(\lambda) < \infty$ for all $\lambda > 0$, then let f^* denote symmetric decreasing rearrangement of f, that is, the unique lower semicontinuous function such that f^* is radially symmetric and nonincreasing with $E_{f^*} = E_f$.

Given a function $F : \mathbb{R}^3 \to \mathbb{R}$, denote its third-order difference by

$$\Delta_3(F; a, b, c, d, e, f) = F(b, d, f) - F(a, d, f) - F(b, c, f) - F(b, d, e) + F(b, c, e) + F(a, d, e) + F(a, c, f) - F(a, c, e).$$

Theorem 2.12. *Let $F : \mathbb{R}^3 \to \mathbb{R}$ be continuous and satisfy*

$$F(0, 0, 0) = F(x, 0, 0) = F(0, y, 0) = F(0, 0, z) = 0, \tag{2.15}$$

along with

$$F(R) := \Delta_3(F; a, b, c, d, e, f) \geq 0 \tag{2.16}$$

for all rectangles $R = \{(x, y, z) : a \leq x \leq b, c \leq y \leq d, e \leq z \leq f\}$.

Then, for any non-negative measurable functions f, g, h on $\mathbb{R}^d$ with finite distribution functions,

$$\iint F(f(s), g(t), h(s+t))dsdt \leq \iint F(f^*(s), g^*(t), h^*(s+t))dsdt. \tag{2.17}$$

Condition (2.15) is simply to ensure that all integrals in the following proof are finite. If $\mathbb{R}$ is replaced with a finite measure space, this condition pay be dropped.

Proof. For this proof, we use the notation

$$I(f,g,h) := \iint F(f(s), g(t), h(s+t))dsdt.$$

By pp.64-68 of [29], we may extend $F(R)$ from a measure on rectangles to a Borel measure on $\mathbb{R}^3$, also denoted by F, provided that F is additive.[1] Here, F is additive if $F(R_1 \cup R_2) = F(R_1) + F(R_2)$ for any non-overlapping rectangles R_1 and R_2. For $F(R_1 \cup R_2)$ to be pre-defined, R_1 and R_2 must have an overlapping face; without loss of generality, assume this face is parallel to the yz-plane. Thus, $R_1 = \{(x,y,z) : a_0 \le x \le a_1, c \le y \le d, e \le z \le f\}$ and $R_2 = \{(x,y,z) : a_1 \le x \le a_2, c \le y \le d, e \le z \le f\}$. By definition of $F(R)$,

$$\begin{aligned} F(R_1) + F(R_2) &= F(a_1,d,f) - F(a_0,d,f) - F(a_1,c,f) - F(a_1,d,e) \\ &\quad + F(a_1,c,e) + F(a_0,d,e) + F(a_0,c,f) - F(a_0,c,e) \\ &\quad + F(a_2,d,f) - F(a_1,d,f) - F(a_2,c,f) - F(a_2,d,e) \\ &\quad + F(a_2,c,e) + F(a_1,d,e) + F(a_1,c,f) - F(a_1,c,e) \\ &= F(a_2,d,f) - F(a_0,d,f) - F(a_2,c,f) - F(a_2,d,e) \\ &\quad + F(a_2,c,e) + F(a_0,d,e) + F(a_0,c,f) - F(a_0,c,e) = F(R_1 \cup R_2). \end{aligned}$$

Let

$$R_{xyz} = \{(\alpha,\beta,\gamma) : 0 \le \alpha \le x, 0 \le \beta \le y, 0 \le \gamma \le z\}$$

be a rectangle with characteristic function

$$\chi_{xyz}(\alpha,\beta,\gamma) = \Phi_{\alpha\beta\gamma}(x,y,z).$$

Then, by (2.15), we have

$$\begin{aligned} \int \chi_{xyz}(\alpha,\beta,\gamma)dF(\alpha,\beta,\gamma) &= F(R_{xyz}) \\ &= F(x,y,z) - F(x,y,0) - F(x,0,z) - F(0,y,z). \end{aligned}$$

[1]The book of Saks [29] proves that F extends to a Borel measure in a similar way that one typically proves volume of rectangles extends to Lebesgue measure. It works by constructing an outer measure F^* in the typical fashion, where $F^*(E)$ is the infimum of $\sum F(R_i)$ for countable collections of rectangles R_i which cover E, and showing that F^* and F agree on rectangles.

Alternatively, one may prove our rearrangement lemma by first assuming that $F \in C^3(\mathbb{R}^3)$, so $dF = F_{xyz}dxdydz$ is well-defined. The third-order condition is used to obtain positivity of the involved integrals. Then, one may extend the result to continuous F by a standard approximation argument which takes F to be the uniform limit of C^3 functions.

Now we substitute $x = f(s), y = g(t), h(s+t)$ and integrate both sides of the above to obtain

$$I(f,g,h) = \iint \left[\int \Phi_{\alpha\beta\gamma}(f(s),g(t),h(s+t))dF(\alpha,\beta,\gamma) \right] dsdt \\ + \iint F(f(s),g(t),0)dsdt + \iint F(f(s),0,h(s+t))dsdt \\ + \iint F(0,g(s),h(s+t))dsdt \quad (2.18)$$

The $\iint F(f(s),g(t),0)dsdt$ term is invariant under symmetrization of f and g since f and g appear as functions of independent variables. The two following terms may be dealt with similarly after a change of variables. Thus it suffices to show that the first term of (2.18) is nondecreasing under rearrangement. By Fubini's theorem,

$$\iint \left[\int \Phi_{\alpha\beta\gamma}(f(s),g(t),h(s+t))dF(\alpha,\beta,\gamma) \right] dsdt = \int J(f,g,h)dF(\alpha,\beta,\gamma).$$

where

$$J(f,g,h) := \iint \Phi_{\alpha\beta\gamma}(f(s),g(t),h(s+t))dsdt.$$

Therefore, using the hypothesis that F is a nonnegative measure, it suffices to show

$$J(f,g,h) \leq J(f^*,g^*,h^*). \quad (2.19)$$

By the steps above, we have in fact shown (2.19) to be equivalent to (2.17). However, note that (2.19) is a statement independent of our choice of F. In the case that $F(x,y,z) = xyz$, then (2.17) is the classical Riesz rearrangement inequality, a previously proven theorem. Hence by this series of equivalences, we have proven our theorem for any F.

□

We conclude this section with the following remark. One may show by example that the third-order condition which is found as a hypothesis in the rearrangement inequality is necessary. To see this, suppose that there exist $a_1 \leq a_2, b_1 \leq b_2, c_1 \leq c_2$ such that $F(R) < 0$, where $R = \{(x,y,z) : a_1 \leq x \leq a_2, b_1 \leq y \leq b_2, c_1 \leq z \leq c_2\}$.

Let $\chi_{[s,t]}$ denote the indicator function of the interval $[s,t]$ and let $f = a_1\chi_{[-5/2,5/2]} + (a_2 - a_1)\chi_{[1/2,3/2]}$, $g = b_1\chi_{[-5/2,5/2]} + (b_2 - b_1)\chi_{[1/2,3/2]}$, and $h = c_1\chi_{[-5,5]} + (c_2 - c_1)\chi_{[-1,1]}$. Denoting $LHS = \iint F(f(s), g(t), h(s+t))dsdt$ and $RHS = \iint F(f^*(s), g^*(t), h^*(s+t))dsdt$, then one may compute

$$\begin{aligned} LHS = F(a_2, b_2, c_1) &+ 2[F(a_2, b_1, c_2) + F(a_2, b_1, c_1) + F(a_1, b_2, c_2) + F(a_1, b_2, c_1)] \\ &+ 5F(a_1, b_1, c_2) + 11F(a_1, b_1, c_1) + F(a_1, 0, c_2) + F(0, b_1, c_2) \\ &+ 5[F(a_2, 0, c_1) + F(0, b_2, c_1)] + 19[F(a_1, 0, c_1) + F(0, b_1, c_1)] \end{aligned}$$

and

$$\begin{aligned} RHS = F(a_2, b_2, c_2) &+ F(a_2, b_1, c_2) + 3F(a_2, b_1, c_1) + F(a_1, b_2, c_2) + 3F(a_1, b_2, c_1) \\ &+ 6F(a_1, b_1, c_2) + 10F(a_1, b_1, c_1) + F(a_1, 0, c_2) + F(0, b_1, c_2) \\ &+ 5[F(a_2, 0, c_1) + F(0, b_2, c_1)] + 19[F(a_1, 0, c_1) + F(0, b_1, c_1)] \end{aligned}$$

Thus, $RHS - LHS = F(R) < 0$.

2.4 The Scales Argument

Let $f, g, h : \mathbb{R}^d \to \mathbb{R}$ and write $f = \sum_{j\in\mathbb{Z}} 2^j F_j$, where $1_{\mathcal{F}_j} \leq |F_j| < 2 \cdot 1_{\mathcal{F}_j}$ and the $\mathcal{F}_j$ are disjoint subsets of $\mathbb{R}^d$. Similarly, we write $g = \sum_{k\in\mathbb{Z}} 2^k G_k$ and $h = \sum_{l\in\mathbb{Z}} 2^l H_l$ with associated sets $\mathcal{G}_k$ and $\mathcal{H}_l$, respectively.

For this section we introduce the following notation. If $B : \mathbb{R}^3_+ \to \mathbb{R}_+$ is measurable, then

$$I_B(f, g, h) := \iint B(f(y), g(x-y), h(x))dxdy.$$

B is an HBL function is and only if $I_B(f, g, h) \lesssim B(\int f, \int g, \int h)$.

Proposition 2.13. *Let $P_i(a, b, c) = a^{1/p_i} b^{1/q_i} c^{1/r_i}$, where $p_i, q_i, r_i \in (1, \infty)$ and $1/p_i + 1/q_i + 1/r_i = 2$. Let $B = \rho(P_1, ..., P_n)$ where*

$$\rho(\lambda_1 y_1, ..., \lambda_n y_n) \leq C \max_i \lambda_i \rho(y_1, ..., y_n) \tag{2.20}$$

and

$$\rho(\vec{y_1}) + \rho(\vec{y_2}) \leq \rho(\vec{y_1} + \vec{y_2}). \tag{2.21}$$

Then there exist positive constants δ_0, c_0, C_0 and positive functions θ, Θ such that

$$\lim_{t\to\infty}\theta(t)=0 \qquad \lim_{\delta\to 0}\Theta(\delta)=0$$

with the following properties. Let $0<\delta\leq\delta_0$ *Let* $f,g,h:\mathbb{R}^d\to[0,\infty)$ *be integrable functions with* $\int f=\alpha,\int g=\beta,\int h=\gamma$ *and*

$$I_B(f,g,h)\geq(1-\delta)AB(\alpha,\beta,\gamma),$$

where A *is the optimal constant in the reverse inequality. Then there exist* $k,k',k''\in\mathbb{Z}$ *such that*

$$2^k|\mathcal{F}_k|\geq c_0$$

$$\sum_{|j-k|\geq m}2^j|\mathcal{F}_j|\leq\theta(m)+\Theta(\delta)$$

with the analogous properties for g *(with* k' *in place of* k*) and* h *(with* k'' *in place of* k*). Lastly, we have*

$$|k-k'|+|k-k''|\leq C_0.$$

Remark 2.14. *It is implicit in the statement of this theorem that* B *is an HBL function. This may be established by using* (2.20) *to prove* (2.7)*. We also note that* (2.20) *is precisely the condition for* ρ *to be an HBL function in the case that each of the* L_j *is the identity map.*

Proof. Let $\eta>0$ be a small parameter and define $S=\{j\in\mathbb{Z}:2^j|\mathcal{F}_j|>\eta\}$. Let $\overline{f}=\sum_{j\in S}2^jF_j$. Note that $|S|\leq C\eta^{-1}$ by Chebyshev's inequality.

Fix $1\leq i\leq n$ and write $p=p_i,q=q_i,r=r_i$. Choose $\tilde{p}>p,\tilde{q}>q,\tilde{r}>r$ with $\frac{1}{\tilde{p}}+\frac{1}{\tilde{q}}+\frac{1}{\tilde{r}}=1$. Then, taking advantage of the disjointness of the $\mathcal{F}_j$, we have

$$\begin{aligned}
||f^{1/p} - \overline{f}^{1/p}||^{\tilde{p}}_{L^{p,\tilde{p}}} &= ||\sum_{j\notin S} 2^{j/p} F_j^{1/p}||^{\tilde{p}}_{L^{p,\tilde{p}}} \\
&\asymp \sum_{j\notin S} (2^{j/p}|\mathcal{F}_j|^{1/p})^{\tilde{p}} \\
&\leq \max_{j\notin S} (2^{j/p}|\mathcal{F}_j|^{1/p})^{\tilde{p}-p} \sum_{j\notin S} (2^{j/p}|\mathcal{F}_j|^{1/p})^{p} \\
&\leq \eta^{\frac{\tilde{p}-p}{p}} \sum_{j\notin S} (2^{j/p}|\mathcal{F}_j|^{1/p})^{p} \\
&\leq C\eta^{\frac{\tilde{p}-p}{p}} ||f^{1/p} - \overline{f}^{1/p}||^{p}_{L^p}.
\end{aligned}$$

Now define $S(\eta) = S \times \mathbb{Z} \times \mathbb{Z}$. Taking advantage of the classical inequality

$$\langle f * g, h\rangle \leq C||f||_{L^{p,\tilde{p}}}||g||_{L^q}||h||_{L^r},$$

we see that

$$\begin{aligned}
I_{P_i}(f - \overline{f}, g, h) &= \sum_{S(\eta)} 2^{j/p_i + k/q_i + l/r_i} \langle F_j^{1/p_i} * G_k^{1/q_i}, H_l^{1/r_i}\rangle \\
&\leq C||f^{1/p} - \overline{f}^{1/p}||_{L^p} \\
&\leq C\eta^{\gamma_i},
\end{aligned}$$

where $\gamma_j = \frac{\tilde{p_i} - p_i}{p_i \tilde{p_i}} > 0$.

By disjointness of supports of f and $\overline{f}$,

$$I_B(f, g, h) = I_B(\overline{f}, g, h) + I_B(f - \overline{f}, g, h)$$

By Theorem 2.1,

$$\iint \rho(f_1(x,y), ..., f_n(x,y))dxdy \leq C\rho\left(\int f_1, ..., \int f_n\right).$$

Thus,

$$\begin{aligned}
I_B(f - \overline{f}, g, h) &\leq \rho\left[I_{P_1}(f - \overline{f}, g, h), ..., I_{P_n}(f - \overline{f}, g, h)\right] \\
&\leq C\rho(C_1\eta^{\gamma_1}, ..., C_n\eta^{\gamma_n}) \\
&\leq C\eta^{\min \gamma_i}
\end{aligned}$$

and

$$I_B(f-\overline{f},g,h)\le C\eta^\gamma \tag{2.22}$$

for some fixed $\gamma>0$.

As $\eta\to 0$, the left hand side of (2.22) approaches 0. However, we are given that f is a near-maximizer of this integral, so $\overline{f}\ne 0$ and $S\ne\emptyset$. This establishes our first conclusion.

For our next conclusions, we will find an upper bound on the diameter of S,

$$M=\max_{j,j'\in S}|j-j'|.$$

Let N be a large positive integer. Then there exist integers $I^\flat<I^\sharp$ such that $S\cap(-\infty,I^\flat]\ne\emptyset, S\cap[I^\sharp,\infty)\ne\emptyset, S\cap(I^\flat,I^\sharp)=\emptyset, I^\sharp-I^\flat\ge M/(2N|S|)$, and, denoting $f_0=\sum_{I^\flat<j<I^\sharp}2^jF_j$,

$$\int|f_0|\le N^{-1}\int|f-\overline{f}|\le CN^{-1}\eta^c.$$

Additionally, we may take $I^\sharp-I^\flat$ to be divisible by 2. Now define

$$f^\sharp=\sum_{j\ge I^\sharp}2^jF_j,\qquad f^\flat=\sum_{j\le I^\flat}2^jF_j$$

so that $f=f^0+f^\sharp+f^\flat$. Next, let $I=(I^\sharp+I^\flat)/2$ and define

$$g^\sharp=\sum_{k\ge I}2^kG_k,\qquad h^\sharp=\sum_{l\ge I}2^lH_l,$$

and $g^\flat=g-g^\sharp, h^\flat=h-h^\sharp$. We will shortly be analyzing the expression

$$\langle (f-f^0)^{1/p}*g^{1/q},h^{1/r}\rangle=\langle (f^\sharp+f^\flat)^{1/p}*(g^\sharp+g^\flat)^{1/q},(h^\sharp+h^\flat)^{1/r}\rangle \tag{2.23}$$

so let us first prove the following lemma.

Lemma 2.15. *There exist constants $c>0$ and $C<\infty$ such that each of the mixed terms in the expansion of* (2.23) *is* $\le C2^{-c\eta M/N}$.

Note that while (2.23) involves nonlinear expressions, we may expand it in a multilinear fashion since $f^\sharp$ and $f^\flat$ have disjoint supports, hence $(f^\sharp+f^\flat)^{1/p}=(f^\sharp)^{1/p}+(f^\flat)^{1/p}$ and so on. To prove the above lemma, we will make use of the following result from [12].

Lemma 2.16. *Let $p, q, r \in (1, \infty)$ with $1/p + 1/q + 1/r = 2$. There exists $\tau > 0$ and $C < \infty$ such that*

$$\langle 1_{\mathcal{F}} * 1_{\mathcal{G}}, 1_{\mathcal{H}} \rangle \leq C \left[\min_{x,y \in \{|\mathcal{F}|, |\mathcal{G}|, |\mathcal{H}|\}} \frac{x}{y} \right]^{\tau} |\mathcal{F}|^{1/p} |\mathcal{G}|^{1/q} |\mathcal{H}|^{1/r} \tag{2.24}$$

for all measurable subsets $\mathcal{F}, \mathcal{G}, \mathcal{H}$ of $\mathbb{R}$ with finite measure.

Proof of Lemma 2.15. Consider the mixed term $\langle (f^\sharp)^{1/p} * (g^\flat)^{1/q}, (h^\sharp)^{1/r} \rangle$ and let $\mathcal{S}$ be the set of multi-indices (j, k, l) such that $j \geq I^\sharp$ and $k < I$. Let $\epsilon > 0$ and $\mathcal{S}^\dagger \subset \mathcal{S}$ be the set of (j, k, l) such that $2^{j/p} |\mathcal{F}_j|^{1/p} \geq \epsilon, 2^{k/q} |\mathcal{G}_k|^{1/q} \geq \epsilon$, and $2^{l/r} |\mathcal{H}_l|^{1/r} \geq \epsilon$. Note that $|\mathcal{S}^\dagger| \leq C\epsilon^{-3}$, a bound which may be obtained by the same reasoning as for our bound on $|S|$. By (2.22), we have

$$\sum_{S \setminus S^\dagger} 2^{j/p+k/q+l/r} \langle 1_{\mathcal{F}_j} * 1_{\mathcal{G}_k}, 1_{\mathcal{H}_l} \rangle \leq C\epsilon^\gamma. \tag{2.25}$$

If $(j, k, l) \in \mathcal{S}^\dagger$, then $2^{j/p} |\mathcal{F}_j|^{1/p} \leq C$ and $2^{k/q} |\mathcal{G}_k|^{1/q}$. The fact that $(j, k, l) \in \mathcal{S}$ implies

$$j \geq I^\sharp \geq I + \frac{1}{4} M/N |S| \geq I + c\eta M/N,$$

so

$$|\mathcal{F}_j| \leq C 2^{-j} \leq C 2^{-I} 2^{-(i-I)} \leq C 2^{-I} 2^{-c\eta M/N}.$$

Also, since $k \leq I$, we have

$$|\mathcal{G}_k| \geq c 2^{-j} \epsilon^q.$$

Therefore,

$$\frac{|\mathcal{F}_j|}{|\mathcal{G}_k|} \leq C\epsilon^{-q} 2^{-c\eta M/N}$$

and (2.24) implies

$$\sum_{\mathcal{S}^\dagger} 2^{j/p+k/q+l/r} \langle 1_{\mathcal{F}_j} * 1_{\mathcal{G}_k}, 1_{\mathcal{H}_l} \rangle \leq C\epsilon^{-C} 2^{-c\eta M/N}. \tag{2.26}$$

Combining (2.25) with (2.26) and choosing ϵ small enough gives

$$\sum_{\mathcal{S}} 2^{j/p+k/q+l/r} \langle 1_{\mathcal{F}_j} * 1_{\mathcal{G}_k}, 1_{\mathcal{H}_l} \rangle \leq C 2^{-c\eta M/N}.$$

This implies the lemma for both $f^\sharp, g^\flat, h^\sharp$ and $f^\sharp, g^\flat, h^\flat$. All other mixed terms may be dealt with similarly.

□

We now observe a simple corollary to the above lemma:

$$\begin{aligned} I_B(f^\sharp, g^\flat, h^\sharp) &= \iint \rho(P_1(f^\sharp(y), g^\flat(x-y), h^\sharp(x)), ..., P_n(f^\sharp(y), g^\flat(x-y), h^\sharp(x)))dxdy \\ &\leq C\rho(I_{P_1}(f^\sharp, g^\flat, h^\sharp), ..., I_{P_n}(f^\sharp, g^\flat, h^\sharp)) \\ &\leq C\rho(C2^{-c\eta M/N}, ..., C2^{-c\eta M/N}) \\ &\leq C2^{-c\eta M/N} \end{aligned}$$

We are almost ready to complete the proof of Proposition 2.13, but we will need to employ the following lemma. It is proven in [12] in the form where $f \in L^p, g \in L^q, h \in L^r$.

Lemma 2.17. *Let $P(y_1, y_2, y_3) = y_1^{1/p} y_2^{1/q} y_3^{1/r}$, where $1 < p, q, r < \infty$. Let $f^\sharp$, $f^\flat$, $g^\sharp$, $g^\flat, h^\sharp, h^\flat$, and η be as before. Then, there exist constants $c, \gamma > 0$, depending only on p, q, r such that*

$$P\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) + P\left(\int f^\flat, \int g^\flat, \int h^\flat\right) \leq (1 - c\eta^\gamma) P\left(\int f, \int g, \int h\right). \tag{2.27}$$

Now, let A be the optimal constant such that $\iint B(f(y), g(x-y), h(x))dxdy \leq AB(\int f, \int g, \int h)$. We apply Lemma 2.15 and the disjointness of supports for $f^\sharp$, $f^\flat$, f_0 to observe that

$$\begin{aligned} I_B(f, g, h) \leq AB\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) &+ AB\left(\int f^\flat, \int g^\flat, \int h^\flat\right) \\ &+ AB\left(\int f_0, \int g, \int h\right) + C2^{-c\eta M/N}. \end{aligned} \tag{2.28}$$

We deal with the f_0 term as follows:

$$\begin{aligned} B\left(\int f_0, \int g, \int h\right) &= \rho\left[P_1\left(\int f_0, \int g, \int h\right), ..., P_n\left(\int f_0, \int g, \int h\right)\right] \\ &= \rho((CN^{-1}\eta^c)^{1/p_1}\beta^{1/q_1}\gamma^{1/r_1}, ...(CN^{-1}\eta^c)^{1/p_n}\beta^{1/q_n}\gamma^{1/r_n}) \\ &\leq CN^{-1}\eta^c. \end{aligned}$$

Now we analyze the first two terms of (2.28). We begin by using the definition of ρ, along with (2.21) to combine everything into a single term containing just ρ and terms found in Lemma 2.17.

$$\begin{aligned}&B\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) + B\left(\int f^\flat, \int g^\flat, \int h^\flat\right)\\ &\leq \rho\Big[P_1\big(\int f^\sharp, \int g^\sharp, \int h^\sharp\big), ..., P_n\big(\int f^\sharp, \int g^\sharp, \int h^\sharp\big)\Big] + \rho\Big[P_1\big(\int f^\flat, \int g^\flat, \int h^\flat\big), ..., P_n\big(\int f^\flat, \int g^\flat, \int h^\flat\big)\Big]\\ &\leq \rho\Big[P_1\big(\int f^\sharp, \int g^\sharp, \int h^\sharp\big) + P_1\big(\int f^\flat, \int g^\flat, \int h^\flat\big), .., P_n\big(\int f^\sharp, \int g^\sharp, \int h^\sharp\big) + P_n\big(\int f^\flat, \int g^\flat, \int h^\flat\big)\Big].\end{aligned}$$

Next, we apply Lemma 2.17, then use (2.20) before returning B to the expression:

$$\begin{aligned}B\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) &+ B\left(\int f^\flat, \int g^\flat, \int h^\flat\right)\\ &\leq \rho\left[(1 - c_1\eta^{\gamma_1})P_1\left(\int f, \int g, \int h\right), ..., (1 - c_n\eta^{\gamma_n})P_n\left(\int f, \int g, \int h\right)\right]\\ &\leq (1 - c\eta^\gamma)\rho\left[P_1\left(\int f, \int g, \int h\right), ..., P_n\left(\int f, \int g, \int h\right)\right]\\ &= (1 - c\eta^\gamma)B\left(\int f, \int g, \int h\right),\end{aligned}$$

where $\gamma = \min_i \gamma_i$ as before.

In summary, we now have:

$$\begin{aligned}A(1-\delta)B\left(\int f, \int g, \int h\right) \leq I_B(f, g, h) \leq A(1 - c\eta^\gamma)B\left(\int f, \int g, \int h\right)&\\ + CN^{-1}\eta^c + C2^{-c\eta M/N},& \quad (2.29)\end{aligned}$$

the first inequality due to the fact that (f, g, h) is a near-maximizing triple. Thus,

$$2^{-c\eta M/N} \geq c\eta^\gamma - cN^{-1}\eta^c - C\delta \geq c\eta^\gamma - cN^{-1} - C\delta.$$

We now choose N to be the integer closest to a sufficiently small multiple of $\eta^{-\gamma}$ so that

$$2^{-c\eta^{1+\gamma}M} \geq c\eta^\gamma - C\delta,$$

so if C_0 is chosen large enough we have $\eta \geq C_0\delta^{1/\gamma}$ implies $M \leq C\eta^{-1-\gamma}(\log \eta)^{-1}$. This completes the proof of the proposition for f and functions g and h may be addressed similarly.

□

Corollary 2.18. *Let S be a compact subset of $(1, \infty)^3$ and let $\{B_k\}_{k=1}^\infty$ be a sequence of functions satisfying the hypotheses of Proposition 2.13 such that the triples of exponents found in the P_i are each contained in S and such that $\lim_{k\to\infty} B_k$ exists, where the limit is taken pointwise. Then the conclusions of Proposition 2.13 hold with $B = \lim_{k\to\infty} B_k$.*

Proof. All but one of the main steps in the proof of the main proposition involve bounding an integral of B. These steps may be repeated with Fatou's lemma as

$$\iint B(*)dxdy \leq \liminf_{k\to\infty} \iint B_k(*)dxdy,$$

where $*$ represents any appropriate collection of functions and the arguments (such as (f, g, h), or $(f^\flat, g^\flat, h^\flat)$, etc.). We complete the last step of the proof as following, using compactness of S to obtain uniform behavior for the constants γ_k and c_k.

$$\begin{aligned} AB\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) + AB\left(\int f^\flat, \int g^\flat, \int h^\flat\right) =& A \lim_{k\to\infty} B_k\left(\int f^\sharp, \int g^\sharp, \int h^\sharp\right) \\ &+ B_k\left(\int f^\flat, \int g^\flat, \int h^\flat\right) \\ \leq& A \liminf_{k\to\infty} (1 - c_k\eta^{\gamma_k}) B_k\left(\int f, \int g, \int h\right) \\ \leq& A(1 - c\eta^\gamma) B\left(\int f, \int g, \int h\right), \end{aligned}$$

where c_k and γ_k are the appropriate constants corresponding to B_k. □

Example 2.5. The previously mentioned example

$$B(y_1, y_2, y_3) = \int_{-1/6}^{1/6} y_1^{2/3-t/2} y_2^{2/3-t/2} y_3^{2/3+t} dt$$

satisfies the conclusions of Proposition 2.13.

2.5 Existence of Maximizers

Following [12], we introduce the following definitions.

Definition 2.6. Let $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ be continuous such that $\lim_{\rho\to\infty} \theta(\rho) = 0$. Then a function $f \in L^1(\mathbb{R}^d)$ is *normalized with norm α with respect to θ* if $\int f = \alpha$ and

$$\int_{|f(x)|>\rho} |f(x)|dx \leq \theta(\rho) \text{ for all } \rho < \infty$$

$$\int_{|f(x)|<\rho^{-1}} |f(x)|dx \leq \theta(\rho) \text{ for all } \rho < \infty.$$

If $\eta > 0$, then $f \in L^1(\mathbb{R}^d)$ is *η-normalized with respect to θ* if there exists a decomposition $f = g + b$ where g is normalized with respect to θ and $||b||_1 < \eta$.

Under the above definitions, Proposition 2.13 states that any extremizing sequence $\{(f_n, g_n, h_n)\}_{n=1}^{\infty}$ for $\iint B(f(y), g(x-y), h(x))dxdy$ may be dilated such that all f_n, g_n, and h_n are η-normalized with their original norms and with respect to the same θ with $\eta \to 0$ as $n \to \infty$.

While this is trivial in the setting involving L^p norms, here we must reference Lemma 2.4, which says $B(\lambda y_1, \lambda y_2, \lambda y_3) = \lambda^2 B(y_1, y_2, y_3)$. Thus, we obtain the dilation symmetry

$$\iint B(\lambda f(\lambda y), \lambda g(\lambda(x-y)), \lambda h(\lambda x))dxdy = \iint B(f(y), g(x-y), h(y))dxdy. \quad (2.30)$$

One may now take each triple (f_n, g_n, h_n) to be at the same scale by application of the dilation symmetry.

We now begin our proof of Theorem 2.2.

Proof. Let $\{(f_n, g_n, h_n)\}_{n=1}^{\infty}$ be an extremizing sequence satisfying $\int f_n = \alpha, \int g_n = \beta, \int h_n = \gamma$ for all $n \geq 1$. By Theorem 2.12 (and a suitable change of coordinate), we may replace f_n, g_n, h_n with (f_n^*, g_n^*, h_n^*) to obtain another extremizing sequence consisting of functions which are radially symmetric and nonincreasing.

By Proposition 2.13 and (2.30), we may replace the extremizing sequence with one which is η-normalized with respect to a continuous function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$, where $\eta \to 0$ as $n \to \infty$. (The benefit here is that we may use the same θ for all triples in our sequence.) In the sequel, $\{(f_n, g_n, h_n)\}_{n=1}^{\infty}$ will denote the new, normalized, symmetrized sequence. To complete the proof, it suffices to show that each of the sequences $\{f_n\}, \{g_n\}, \{h_n\}$ is precompact.

Let $\epsilon > 0$. For any $\rho < \infty$ and $0 < A < \infty$ we have

$$\int_{|t| \leq A} f_n(t)dt \leq c_d \rho A + \int_{f_n > \rho} f_n.$$

Since f_n is η-normalized with $\eta \to 0$, there exist ρ and N large enough such that $n > N$ implies

$$\int_{f_n > \rho} f_n < \epsilon/2.$$

By choosing A small enough, we have

$$\int_{|t| \leq A} f_n(t)dt < \epsilon \quad (2.31)$$

for sufficiently large n. Now let $0 < B < \infty$. Since symmetric decreasing f_n with $\int f_n = \alpha$ satisfy $f_n(s) \leq c_d \alpha |s|^{-d}$, we have

$$\int_{|t|\geq B} f_n(t) \leq \int_{|t|\geq B} f_n \leq c_d \alpha B^{-d} f_n(t)dt \leq \theta(c_d^{-1}\alpha^{-1}B^d) + o(1),$$

where $o(1) \to 0$ as $n \to \infty$. Since $\theta(\rho) \to 0$ as $\rho \to \infty$, we may take B large enough that

$$\int_{|t|\geq B} f_n(t)dt < \epsilon \tag{2.32}$$

for sufficiently large n. Fixing $0 < A < B < \infty$, we see that the restrictions of f_n to $[A, B]$ are radial symmetric decreasing with $0 \leq f_n(t) \leq c_d \alpha A^{-d}$ so they are precompact in L^1 on $\{t \in \mathbb{R}^d : A \leq |t| \leq B\}$. By (2.31) and (2.32), $\{f_n\}$ is precompact in $L^1(\mathbb{R}^d)$. By the same reasoning, $\{g_n\}$ and $\{h_n\}$ are precompact in $L^1(\mathbb{R}^d)$ as well, which completes the proof.

□

2.6 Non-Gaussian Maximizers

In the classical version of Young's inequality, it is known that maximizers exist when $p_j \in (1, \infty$ and $\sum_j p_j^{-1} = 2$. Furthermore, those maximizers are always triples of Gaussians. In [23], it is shown that for a certain class of functions B, there exist maximizers of

$$\iint B(f(y), g(x-y), h(x))dxdy$$

and that these maximizers are always Gaussians. However, the following proposition shows that our expansion of the class of functions B breaks this pattern.

Proposition 2.19. *Fix $\alpha, \beta, \gamma > 0$. There exists a $B : \mathbb{R}^3 \to \mathbb{R}$ satisfying the hypotheses of Theorem 2.2 such that under the constraints $\int f = \alpha, \int g = \beta, \int h = \gamma$, there exist maximizers of*

$$\iint B(f(y), g(x-y), h(x))dxdy$$

which are not all Gaussians.

The proof of this proposition is based on a simple use of Euler-Lagrange equations, though some aspects are modified to fit our particular setting. Extremizers exist due to results from previous sections and extremizers must also be critical points of the functional $\int B(f(x), g(x-y), h(y))$. However, any critical point must satisfy the Euler-Lagrange equations and it will be clear that no triple of Gaussians does.

To conduct this analysis, let us define a *critical point* as a triplet of L^1 functions (f, g, h) such that for any $j \in C_c^\infty$ with $\int j = 0$,

$$\iint B(f(y)+tj(y), g(x-y), h(x))dxdy = \iint B(f(y), g(x-y), h(x))dxdy + o(|t|)$$

as $t \to 0$ and that the analogous equation holds with perturbations of g and h. The reason we add the restriction that $\int j = 0$ is so that $\int (f+j) = \int f = \alpha$ and $f + j$ satisfies the appropriate constraint. The condition that j is bounded with compact support is to ensure convergence of certain integrals which arise in the proof.

Proof. Let $B(y_1, y_2, y_3) = y_1^{1/p_1} y_2^{1/p_2} y_3^{1/p_3} + y_1^{1/q_1} y_2^{1/q_2} y_3^{1/q_3}$, where

$$\frac{1}{p_i} + \frac{1}{q_i} + \frac{1}{r_i} = 2$$

and $p_i, q_i, r_i \in (1, \infty)$ for $i = 1, 2$, but $(p_1, q_1, r_1) \neq (p_2, q_2, r_2)$. Suppose, to the contrary, that there exist Gaussians f, g, h which are maximizers of $\iint B(f(y), g(x-y), h(x))dxdy$. Then, f, g, h must also form a critical point. Taking the binomial expansion of $(f + tj)^{1/p_1}$, we find

$$\begin{aligned}\iint (f(y)+tj(y))^{1/p_1} g^{1/q_1}(x-y) h^{1/r_1}(x)dxdy = &\iint f^{1/p_1}(y) g^{1/q_1}(x-y) h^{1/r_1}(x)dxdy \\ &+ \frac{t}{p_1} \iint f^{1/p_1 - 1}(y) j(y) g^{1/q_1}(x-y) h^{1/r_1}(x)dxdy \\ &+ O\left(t^2 \iint f^{1/p_1 - 2}(y) j^2(y) g^{1/q_1}(x-y) h^{1/r_1}(x)dxdy\right).\end{aligned}$$

The left hand side is well-defined since f is bounded below by a positive constant on the domain of j. Thus, we may take t small enough that $f + tj > 0$ everywhere. Furthermore, the integrals on the right hand side are convergent since j is bounded with compact support and $1/f$ is bounded on the support of j. In fact, $f^{1/p_1 - 1} j \in L^p$ for all $1 \le p \le \infty$. Thus,

$$\iint f^{1/p_1-1}(y)j(y)g^{1/q_1}(x-y)h^{1/r_1}(x)dxdy + \iint f^{1/p_2-1}(y)j(y)g^{1/q_2}(x-y)h^{1/r_2}(x)dxdy = 0$$

for all bounded j with compact support with $\int j = 0$. This implies that

$$f^{1/p_1-1}(\tilde{g}^{1/q_1} * h^{1/r_1}) + f^{1/p_2-1}(\tilde{g}^{1/q_2} * h^{1/r_2}) = C$$

for some constant C, where $\tilde{g}(x) = g(-x)$. There are now two cases. The first is that neither of the 2 summed terms is constant, in which case each is either a Gaussian or the inverse of a Gaussian and their sum cannot be constant. The second case is that each of the two terms is constant. However, since $(p_1, q_1, r_1) \neq (p_2, q_2, r_2)$, this is impossible to obtain with the same Gaussians for each term. Thus, a triple of Gaussians cannot be a critical point (or a triple of maximizers) for $\iint B(f(y), g(x-y), h(x))dxdy$ under the given constraints. $\square$

Chapter 3

A Sharpened Inequality for Twisted Convolution

3.1 Introduction

Young's convolution inequality, in its optimal form, states that for dimensions $d \geq 1$ and functions $f \in L^p(\mathbb{R}^d), g \in L^q(\mathbb{R}^d)$,

$$||f * g||_{L^r} \leq \mathbf{A}_{\mathbf{p}}^d ||f||_{L^p} ||g||_{L^q}, \tag{3.1}$$

where $p, q, r \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. $\mathbf{A}_{\mathbf{p}}^d = \prod_{j=1}^3 C_{p_j}^d$ is the optimal constant, where $C_p = p^{1/p}/p'^{1/p'}$, and p' is the conjugate exponent of p [4], [9]. For the purpose of this chapter, it is convenient to use the following, related trilinear form:

$$\mathcal{T}(f_1, f_2, f_3) = \iint f_1(x) f_2(y) f_3(x+y) dx dy. \tag{3.2}$$

Through duality, one may rewrite (3.1) as

$$|\mathcal{T}(\mathbf{f})| \leq \mathbf{A}_{\mathbf{p}}^d \prod_{j=1}^3 ||f_j||_{L^{p_j}} \tag{3.3}$$

for all $\mathbf{f} = (f_j \in L^{p_j}(\mathbb{R}^d) : j = 1, 2, 3)$, with $\sum p_j^{-1} = 2$ and $\mathbf{p} = (p_j : j = 1, 2, 3) \in [1, \infty]^3$.

From here on out, we take $p_j \in (1, \infty)$. In [9], Brascamp and Lieb show that the maximizers of (3.3) are precisely the triple of Gaussians $\mathbf{g} = (e^{-\pi p_j' |x|^2} : j = 1, 2, 3)$ and its orbit under the following symmetries.

- $(f_1, f_2, f_3) \mapsto (af_1, bf_2, cf_3)$ for $a, b, c \neq 0$. (Scaling)
- $(f_1, f_2, f_3) \mapsto (M_\xi f_1, M_\xi f_2, M_{-\xi} f_3)$, where $M_\xi f(x) = e^{ix\cdot\xi}$ for $\xi \in \mathbb{R}^d$. (Modulation)
- $(f_1, f_2, f_3) \mapsto (\tau_{v_1} f_1, \tau_{v_2} f_2, \tau_{v_1+v_2} f_3)$, where $\tau_v f(x) = f(x+v)$ for $v \in \mathbb{R}^d$. (Translation)
- $(f_1, f_2, f_3) \mapsto (f_1 \circ \psi, f_2 \circ \psi, f_3 \circ \psi)$, where ψ is an invertible linear map on $\mathbb{R}^d$. (Diagonal Action of the General Linear Group)

Note that these symmetries do not necessarily preserve $|\mathcal{T}(\mathbf{f})|$, but they do preserve $|\Phi(\mathbf{f})|$, where $\Phi(\mathbf{f}) := \frac{\mathcal{T}(\mathbf{f})}{\prod_j ||f_j||_{p_j}}$.

Let $\mathcal{O}_C(\mathbf{f})$ denote the orbit of the triple $\mathbf{f}$ under the above symmetries. Define the distance from $\mathbf{g}$ to $\mathcal{O}_C(\mathbf{f})$ as

$$\operatorname{dist}_{\mathbf{p}}(\mathcal{O}_C(\mathbf{f}), \mathbf{g}) := \inf_{\mathbf{h}\in\mathcal{O}_C(\mathbf{f})} \max_j ||h_j - g_j||_{p_j}. \tag{3.4}$$

Note that the symmetries of an operator preserve the (normalized) distance of a triple from the manifold of maximizers.

Christ [17] proved the following quantitative stability theorem for Young's convolution inequality.

Theorem 3.1. *Let K be a compact subset of $(1,2)^3$. Let $\mathbf{p}$ satisfy $\sum_{j=1}^3 p_j^{-1} = 2$. For each $d \geq 1$, there exists $c > 0$ such that for all $\mathbf{p} \in K$ and all $\mathbf{f} \in L^{\mathbf{p}}(\mathbb{R}^d)$,*

$$|\mathcal{T}(\mathbf{f})| \leq \left(\mathbf{A}_{\mathbf{p}}^d - c\,dist_{\mathbf{p}}(\mathcal{O}_C(\mathbf{f}), \mathbf{g})^2\right) \prod_j ||f_j||_{p_j}. \tag{3.5}$$

One may instead state the above theorem in terms of the distance of a triple $\mathbf{f}$ from the set of all triples of maximizers (that is, $\mathcal{O}_C(\mathbf{g})$), as is done in [17]. However, the distance defined in (4.4) is more useful for analogy with our current analysis.

It is also shown that the conclusion of Theorem 3.1 is true for $\mathbf{p} \in (1,2]^3$ provided one does not require the same c for all $\mathbf{p}$ in a region. (However it is not known if this uniformity fails.) Furthermore, the conclusion in this particular quantitative form is false whenever any $p_j = 1$ or $p_j > 2$.

The purpose of this chapter is to prove a similar quantitative stability result for twisted convolution. Let $t \geq 0$ be a parameter and let $f_j \in L^{p_j}(\mathbb{R}^{2d})$, where $\mathbb{R}^{2d}$ is viewed as $\mathbb{R}^d \times \mathbb{R}^d = \{(x', x''') : x', x'' \in \mathbb{R}^d\}$. Define the trilinear twisted convolution form with parameter t as

$$\mathcal{T}_t(\mathbf{f}) := \iint f_1(x)f_2(y)f_3(x+y)e^{it\sigma(x,y)}dxdy, \tag{3.6}$$

where $\sigma(x,y) = x' \cdot y'' - x'' \cdot y'$ is the symplectic form. It is often useful to write $\sigma(x,y) = x^tJy$, where J is the matrix

$$J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}, \tag{3.7}$$

and I_d is the $d \times d$ identity matrix.

When $t = 0$, (3.6) becomes the trilinear form representing convolution. When $t \neq 0$, it is obvious through the inequality

$$|\mathcal{T}_t(\mathbf{f})| \leq \mathcal{T}(|\mathbf{f}|) \tag{3.8}$$

that $\mathcal{T}_t$ is bounded for any triple $\mathbf{p}$ of exponents for which $\mathcal{T}$ is bounded. It is also known that for $t \neq 0$, $\mathcal{T}(\mathbf{f},t)$ is also bounded for $(p_1,p_2,p_3) = (2,2,2)$ and the full range of exponents implied by interpolation (see Chapter XII.4 of [31], for instance). However, the particular conclusion we desire is false in the case $\sum_j p_j^{-1} \neq 2$ since $\mathcal{T}_0 = \mathcal{T}$ is unbounded.

By (3.8), it is easy to see that $\mathcal{T}_t$ has norm at most $A_{\mathbf{p}}^{2d}$, the optimal constant for Young's convolution inequality. Furthermore, the optimal constant may be seen to equal $A_{\mathbf{p}}^{2d}$ by taking a triple of Gaussians which optimize Young's inequality and dilating them to concentrate at the origin so the oscillation of the twisting factor has negligible effect. However, no extremizers of $\mathcal{T}_t$ exist for fixed $t \neq 0$. [24]

One challenge to dealing with the above form directly arises because the symmetry group of $\mathcal{T}$ contains the general linear group $Gl(2d)$, while $\mathcal{T}_t$ does not; the only linear transformations which preserve σ are the symplectomorphisms. To avoid this issue, it helps to introduce the following trilinear form:

$$\mathcal{T}_A(\mathbf{f}) := \iint f_1(x)f_2(y)f_3(x+y)e^{it\sigma(Ax,Ay)}dxdy, \tag{3.9}$$

where $A : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is an arbitrary linear map. Replacing x with Lx and y with Ly for an invertible matrix L sends A to $A \circ L$, and the functional remains of the form (3.9). Boundedness properties of $\mathcal{T}_A$ follow directly from those of $\mathcal{T}_t$ and a change of coordinates.

The symmetries of $\mathcal{T}_A$ are similar to the those of $\mathcal{T}$ with some slight modifications, though they reduce to the symmetries of $\mathcal{T}(\mathbf{f})$ when $A = 0$. Here, the symmetries preserve $|\Phi(\mathbf{f},A)|$, where $\Phi(\mathbf{f},A) = \frac{\mathcal{T}(\mathbf{f},A)}{\prod_j ||f_j||_{p_j}}$.

- $(f_1, f_2, f_3, A) \mapsto (af_1, bf_2, cf_3, A)$, where $a, b, c \in \mathbb{C}$. (Scaling)
- $(f_1, f_2, f_3, A) \mapsto (M_\xi f_1, M_\xi f_2, M_{-\xi} f_3, A)$. (Modulation)
- $(f_1, f_2, f_3, A) \mapsto (M_{A^T JAv_2}\tau_{v_1} f_1, M_{-A^T JAv_1}\tau_{v_2} f_2, \tau_{v_1+v_2} f_3, A)$, where A^T represents the transpose of the matrix A. (Translation/ Modulation Mix)
- $(f_1, f_2, f_3, A) \mapsto \mathcal{T}(f_1 \circ \psi, f_2 \circ \psi, f_3 \circ \psi, A \circ \psi)$, where $\psi \in Gl(d)$. (Diagonal Action of the General Linear Group)

Note that only the last of these symmetries alters A.

Let $\mathcal{O}_{TC}(\mathbf{f}, A)$ denote the orbit of $(\mathbf{f}, A)$ under the above symmetries.

Now, it is less obvious how to represent the distance of A from the zero transformation than it was when our parameter was just a real number t. One may naively suggest that $||A||$ will play a role, but this approach ignores the role of the symplectic group. The real symplectic group $Sp(2d)$ is defined as the set of invertible $(2d) \times (2d)$ matrices S such that $S^T JS = J$. Equivalently, $Sp(2d)$ may be viewed as the set of coordinate changes which preserve σ. Under this view, we see that $\sigma(Ax, Ay) = \sigma(SAx, SAy)$ for any $S \in Sp(2d)$. Thus, replacing A with $S \circ A$ should not change our distance.

With this in mind, define the distance from $\mathcal{O}_{TC}(\mathbf{f}, A)$ to $(\mathbf{g}, 0)$ by

$$\text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, A), (\mathbf{g}, 0))^2 := \inf_{(\mathbf{h}, M) \in \mathcal{O}_{TC}(\mathbf{f}, A)} \left[\max_j ||h_j - g_j||_{p_j}^2 + ||M^T JM||^2 \right] \quad (3.10)$$

A useful fact in analyzing this distance is that $\inf_{S \in Sp(2d)} ||S \circ A||^2 = ||A^T JA||$. (See Lemma 10.1 of [16].) Define $||\mathbf{f}||_{\mathbf{p}} = \max_j ||f_j||_{p_j}$. We now state our main theorem.

Theorem 3.2. *Let K be a compact subset of $(1, 2)^3$. For each $d \geq 1$, there exists $c > 0$ such that for all $\mathbf{p} \in K$ with $\sum_{j=1}^3 p_j^{-1} = 2$, $\mathbf{f} \in L^{\mathbf{p}}(\mathbb{R}^{2d})$, and $(2d) \times (2d)$ matrices A,*

$$|\mathcal{T}_A(\mathbf{f})| \leq \left(\mathbf{A}_{\mathbf{p}}^{2d} - c\,dist_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, A), (\mathbf{g}, 0))^2 \right) \prod_j ||f_j||_{p_j}. \quad (3.11)$$

By setting $A = t^{1/2} I_{2d}$ in Theorem 3.2 (where I_{2d} is the $(2d) \times (2d)$ identity matrix), one obtains the following corollary. However, one is cautioned that the orbit in this expression refers to the symmetries of $\mathcal{T}_A$, not those of $\mathcal{T}_t$.

Corollary 3.3. *Let K be a compact subset of $(1,2)^3$. For each $d \geq 1$, there exists $c > 0$ such that for all $\mathbf{p} \in K$ with $\sum_{j=1}^{3} p_j^{-1} = 2$, $\mathbf{f} \in L^{\mathbf{p}}(\mathbb{R}^{2d})$, and $|t| \leq 1$,*

$$|\mathcal{T}_t(\mathbf{f})| \leq \left(\mathbf{A}_{\mathbf{p}}^{2d} - c\,dist_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, t^{1/2}I_{2d}), (\mathbf{g}, 0))^2\right) \prod_j ||f_j||_{p_j}. \tag{3.12}$$

The reason one uses $t^{1/2}I_{2d}$ rather than tI_{2d} is so the $||M^T JM||^2$ term appearing in (3.10) is proportional to t^2, rather than t^4. An alternative form of Corollary 3.3 states the function $\epsilon(\delta)$ in Theorem 3.4 may be taken to be $C\sqrt{\delta}$ for some $C > 0$.

The methods in this chapter follow the general approach found in [17] and [7] in which one takes a Taylor-like expansion of the given operator and diagonalizes the resulting quadratic form.

We will often use C or c to denote an arbitrary constant in $(0, \infty)$ which may change from line to line but always be independent of functions found in the equation.

3.2 Reduction to Perturbative Case

Our argument centers around an expansion of $\mathcal{T}(\mathbf{f}, A)$ which requires a reduction to small perturbations. To this end, the following result from [16] is essential.

Theorem 3.4. *Let $d \geq 1$. Let K be a compact subset of $(1,2)^3$ for which each $\mathbf{p} \in K$ satisfies $\sum_{j=1}^{3} p_j^{-1} = 2$. Then, there exists a function $\delta \mapsto \epsilon(\delta)$ (depending only on K and d) satisfying $\lim_{\delta \to 0} \epsilon(\delta) = 0$ with the following property. Let $\mathbf{f} \in L^{\mathbf{p}}(\mathbb{R}^{2d})$ and suppose that $||f_j||_{p_j} \neq 0$ for each $1 \leq j \leq 3$. Let $\delta \in (0,1)$ and suppose that $|\mathcal{T}(\mathbf{f}, t)| \geq (1-\delta)\mathbf{A}_{\mathbf{p}}^{2d} \prod_j ||f_j||_{p_j}$. Then there exist $S \in Sp(2d)$ and a triple of Gaussians $\mathbf{G} = (G_1, G_2, G_3)$ such that $G_j^{\natural} = G_j \circ S$ satisfy*

$$||f_j - G_j^{\natural}||_{p_j} < \epsilon(\delta)||f_j||_{p_j} \tag{3.13}$$

for $1 \leq j \leq 3$ and

$$G_j(x) = c_j e^{\pi p_j' |L(x-a_j)|^2} e^{ix\cdot v} e^{it\sigma(\tilde{a_j}, x)} \tag{3.14}$$

where $v \in \mathbb{R}^{2d}, 0 \neq c_j \in \mathbb{C}, a_1 + a_2 + a_3 = 0, \tilde{a_3} = 0, \tilde{a_1} = a_2, \tilde{a_2} = a_1, L \in Gl(2d)$, and

$$|t| \cdot ||L^{-1}||^2 \leq \epsilon(\delta). \tag{3.15}$$

Here is a rephrasing of Theorem 3.4.

Theorem 3.5. *Let $d \geq 1$. Let K be a compact subset of $(1,2)^3$ for which each $\mathbf{p} \in K$ satisfies $\sum_{j=1}^{3} p_j^{-1} = 2$. Then, there exists a function $\delta \mapsto \epsilon(\delta)$ (depending only on*

K and d) satisfying $\lim_{\delta\to 0}\epsilon(\delta)=0$ *with the following property. Let* $\mathbf{f}\in L^{\mathbf{p}}(\mathbb{R}^{2d})$ *and suppose that* $||f_j||_{p_j}\neq 0$ *for each* $1\leq j\leq 3$. *Let* $\delta\in(0,1)$ *and suppose that* $|\mathcal{T}_A(\mathbf{f})|\geq(1-\delta)\mathbf{A}_{\mathbf{p}}^{2d}\prod_j||f_j||_{p_j}$. *Then,*

$$dist_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f},A),(\mathbf{g},0))<\epsilon(\delta) \tag{3.16}$$

Proof of Theorem 3.4 ⇒ Theorem 3.5. By a standard approximation argument, it suffices to prove Theorem 3.5 for invertible maps A, as each noninvertible map is arbitrarily close to an invertible map.

Suppose that $|\mathcal{T}(\mathbf{f},A)|\geq(1-\delta)\mathbf{A}_{\mathbf{p}}^{2d}\prod_j||f_j||_{p_j}$. Then invoking the symmetry of diagonal action of the general linear group,

$$|\mathcal{T}(\mathbf{f}\circ A^{-1},I_{2d})|\geq(1-\delta)\mathbf{A}_{\mathbf{p}}^{2d}\prod_j||f_j\circ A^{-1}||_{p_j}, \tag{3.17}$$

where $\mathbf{f}\circ A^{-1}=(f_j\circ A^{-1}:j=1,2,3)$.

Applying Theorem 3.4 under the case $t=1$, there exists $S_0\in Sp(2d)$ and a triple of Gaussians $\mathbf{G}=(G_1,G_2,G_3)$ such that

$$||f_j\circ A^{-1}-G_j\circ S_0||_{p_j}<\epsilon(\delta)||f_j\circ A^{-1}||_{p_j} \tag{3.18}$$

for $1\leq j\leq 3$ and

$$G_j(x)=c_je^{\pi p_j'|L(x-a_j)|^2}e^{ix\cdot v}e^{it\sigma(\tilde{a_j},x)} \tag{3.19}$$

where $v\in\mathbb{R}^{2d}, 0\neq c_j\in\mathbb{C}, a_1+a_2+a_3=0, \tilde{a_3}=0, \tilde{a_1}=a_2, \tilde{a_2}=a_1, L\in Gl(2d)$, and

$$||L^{-1}||^2\leq\epsilon(\delta). \tag{3.20}$$

By a combination of translations, modulations, scalings, and compositions with invertible linear maps, (3.18) becomes

$$||h_j-g_j||_{p_j}<\epsilon(\delta)||h_j||_{p_j}, \tag{3.21}$$

where h_j is $f_j\circ A^{-1}$ composed with said operations.

Since $\mathbf{G}$ was the composition of $\mathbf{g}$ with the stated symmetries of $\mathcal{T}_A$, we see that $\mathbf{h}$ is obtained by the composition of $\mathbf{f}\circ A^{-1}$ with symmetries of $\mathcal{T}_A$ by the following reasoning. Three of these symmetries (scaling, modulation, and the diagonal action of the general linear group) may trivially be inverted by symmetries of the same form. To address the inversion of the translation/modulation mix, one observes that $\tau_{w_j}M_{B^TJB\tilde{w_j}}f=e^{iB^TJB\tilde{w_j}\cdot w_j}M_{B^TJB\tilde{w_j}}\tau_{w_j}f$ for matrices B and vectors w_j. Hence, $\mathbf{h}$ is obtained from $\mathbf{f}\circ A^{-1}$ through the inverses of the symmetries applied initially to $\mathbf{g}$ to obtain $\mathbf{G}$ but with an additional scaling symmetry.

The only above symmetry which changes the matrix B in $\mathcal{T}_B$ is the diagonal action of the general linear group. Following the use of this symmetry above, one obtains from (3.17) that $(\mathbf{h}, S_0^{-1} \circ L^{-1}) \in \mathcal{O}_{TC}(\mathbf{f}, A)$.

We now see that

$$\begin{aligned} \text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, A), (\mathbf{g}, 0))^2 &\leq \max_j ||h_j - g_j||_{p_j}^2 + \inf_{S \in Sp(2d)} ||S \circ S_0^{-1} \circ L^{-1}||^4 \\ &\leq \epsilon(\delta)^2 + ||S_0 S_0^{-1} \circ L^{-1}||^4 \\ &\leq \epsilon(\delta)^2 + ||L^{-1}||^4 \leq 2\epsilon(\delta)^2. \end{aligned}$$

□

As a corollary to Theorem 3.5, it suffices to prove Theorem 3.2 in the case in which $\text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, A), (\mathbf{g}, 0)) < \delta_0$ for some $\delta_0 > 0$. Theorem 3.5 guarantees that there are no sequences of $(\mathbf{f}_n, A_n)$ at distance greater than δ_0 such that $\mathcal{T}_{A_n}(\mathbf{f}_n)/(\prod_j ||f_{n,j}||_{p_j})$ converges to $A_{\mathbf{p}}^{2d}$. Thus, for $(\mathbf{f}, A)$ at distance at least δ_0, $\mathcal{T}_A(\mathbf{f})$ must have a maximum strictly less than $A_{\mathbf{p}}^{2d}$. While $||A^T J A|| \to \infty$ for an appropriate sequence of matrices A, $\text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{f}, A), (\mathbf{g}, 0))$ remains bounded above as the symmetries of $\mathcal{T}_A$ ensure there exists $(\mathbf{h}, M) \in \mathcal{O}_{TC}(\mathbf{f}, A)$ with $||M^T J M|| \leq 1$. Therefore, the conclusion of Theorem 3.2 holds for distances greater than δ_0.

3.3 Treating Some Terms of the Expansion

In this section, we consider $\mathcal{T}_A(\mathbf{g} + \mathbf{f})$, where A is a $(2d) \times (2d)$ matrix, $\mathbf{g} = (g_j = e^{-\pi p_j' |x|^2} : j = 1, 2, 3)$ and $\mathbf{f} \in L^{\mathbf{p}}(\mathbb{R}^{2d})$ are small perturbations. (This change in notation of $\mathbf{f}$ from functions close to $\mathbf{g}$ to the differences will continue for the remainder of the chapter.) As in [17], we may assume $\int g_j^{p_j - 1} f_j = 0$ via the scaling symmetry.

In short, we will expand $\mathcal{T}(\mathbf{g} + \mathbf{f}, A) = \mathcal{T}_0(\mathbf{g} + \mathbf{f}) + (\mathcal{T}_A - \mathcal{T}_0)(\mathbf{g} + \mathbf{f})$ and use the multilinearity of $\mathcal{T}_0$ and $\mathcal{T}_A$ to get sixteen terms of eight different types. Before writing out the expansion, we prove a few lemmas about its terms and describe a useful decomposition.

Following [14] and [17], let $\eta > 0$ be a small parameter to be chosen later (see Proposition 3.10). For each $1 \leq j \leq 3$, decompose $f_j = f_{j,\sharp} + f_{j,\flat}$, where

$$f_{j,\sharp} = \begin{cases} f_j(x) \text{ if } |f_j(x)| \leq \eta g_j(x) \\ 0 \text{ otherwise,} \end{cases} \tag{3.22}$$

and $f_{j,\flat} = f_j - f_{j,\sharp}$. The purpose of this decomposition is twofold. First, it is used in the analysis of [17] to analyze the quadratic form in the expansion with L^2

functions. Using the same decomposition allows us to borrow from that analysis in Proposition 3.10, a version of Theorem 3.1 with an additional favorable term. Second, the decomposition is used to reduce to the case of $f_j = f_{j,\sharp}$, which concentrates closer to the origin, allowing for control of the third order term in Lemma 3.7.

Lemma 3.6. $(\mathcal{T}_A - \mathcal{T}_0)(\mathbf{f}) = O(||\mathbf{f}||_{\mathbf{p}}^3)$.

Proof. This claim follows trivially from the uniform boundedness of $\mathcal{T}_A$ and $\mathcal{T}_0$. □

The following lemma represents our main use of the $f_j = f_{j,\sharp} + f_{j,\flat}$ decomposition and the swapping of f_j for $f_{j,\sharp}$ will be justified later.

Lemma 3.7. $(\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp}, f_{2,\sharp}, g_3) = o(||\mathbf{f}||_{\mathbf{p}}^2 + ||A^T JA||^2)$ *with decay rate depending only on* η*.*

Lemma 3.7 also applies to the other two terms of this type.

Note that the trivial bound

$$|(\mathcal{T}_A - \mathcal{T}_0)(h_1, h_2, g_3)| = O\left(||h_1||_{p_1}||h_2||_{p_2}\right) \tag{3.23}$$

is insufficient to deal with the above term directly since it provides a second order control of a term which should heuristically be third order. However, (3.23) still plays a useful role in the proof of Lemma 3.7.

Proof. First, suppose that $||A^T JA||^3 \geq ||f_{1,\sharp}||_{p_1}||f_{2,\sharp}||_{p_2}$. Note that by our reduction to small perturbations in Theorem 3.5, $||A^T JA||$ may be taken small enough that $||A^T JA||^3 \leq ||A^T JA||^2$. By (3.23),

$$(\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp}, f_{2,\sharp}, g_3) \leq C||f_{1,\sharp}||_{p_1}||f_{2,\sharp}||_{p_2} \leq ||A^T JA||^3 = o(||\mathbf{f}||_{\mathbf{p}}^2 + ||A^T JA||^2)$$

and we are done.

So suppose that $||A^T JA||^3 < ||f_{1,\sharp}||_{p_1}||f_{2,\sharp}||_{p_2}$. Now, for $j = 1, 2$, write $f_{j,\sharp} = f_{j,\sharp,\leq M_j} + f_{j,\sharp,>M_j}$, where $f_{j,\sharp,\leq M_j} = f_{j,\sharp}\mathbf{1}_{B(0,M_j)}$ and $f_{j,\sharp,>M_j} = f_{j,\sharp}\mathbf{1}_{B(0,M_j)^c}$. In the above, $\mathbf{1}_E$ refers to the indicator function of the set E, $B(x_0, R)$ refers to the closed ball of radius R centered at x_0, E^c is the complement of the set E, and M_j is chosen so that

$$||f_{j,\sharp,>M_j}||_{p_j} = ||f_{j,\sharp}||_{p_j}^2. \tag{3.24}$$

Note that M_j is dependent on η.

We claim that $M_j \leq C\log(||f_{j,\sharp}||_{p_j}^{-1})$. To see this, observe that for given η and $||f_{j,\sharp}||_{p_j}$ and varying $f_{j,\sharp}$, M_j is maximized when $f_{j,\sharp} = \eta g_j$ on $B(0, M)^c$ and $f_{j,\sharp} = 0$ on $B(0, M)$, where M is the positive real number that leads to the appropriate value of $||f_{j,\sharp}||_{p_j}$. (Here, $M < M_j$ since $||f_{j,\sharp}||_{p_j}$ is small.) It suffices to find an upper bound for M_j in this scenario. We integrate with respect to spherical coordinates to obtain

$$\begin{aligned} ||f_{j,\sharp}||_{p_j}^2 &= ||f_{j,\sharp,>M_j}||_{p_j} \\ &= \int_{S^{d-1}} \left[\int_{M_j}^{\infty} \eta e^{-\pi p_j' r^2} r^{2d-1} dr \right] d\sigma(\theta) \\ &= C_d \eta \int_{M_j}^{\infty} \eta e^{-\pi p_j' r^2} r^{2d-1} dr \\ &= O(M_j^{2d-2} e^{-\pi p_j' M_j^2}). \end{aligned}$$

Thus, $||f_{j,\sharp}||_{p_j} \le Ce^{-M_j}$, proving our claim.

Expand

$$\begin{aligned} (\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp}, f_{2,\sharp}, g_3) = (\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp,>M_1}, f_{2,\sharp,>M_2}, g_3) + (\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp,>M_1}, f_{2,\sharp,\le M_2}, g_3) \\ + (\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp,\le M_1}, f_{2,\sharp,>M_2}, g_3) + (\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp,\le M_1}, f_{2,\sharp,\le M_2}, g_3). \end{aligned}$$

The first three of these terms may be treated by combining the trivial bound (3.23) with (3.24).

Let $R = B(0, M_1) \times B(0, M_2) \subset \mathbb{R}^{2d} \times \mathbb{R}^{2d}$. The absolute value of the remaining term is

$$\begin{aligned} |(\mathcal{T}_A - \mathcal{T}_0)(f_{1,\sharp}, f_{2,\sharp}, g_3)| &\le \iint_R |f_{1,\sharp}(x)| \cdot |f_{2,\sharp}(y)| \cdot g_3(x+y) \cdot |\sigma(Ax, Ay)| dx dy \\ &\le C ||f_{1,\sharp}||_{p_1} ||f_{2,\sharp}||_{p_2} ||g_3||_{p_3} ||A^T J A|| M_1 M_2 \\ &\le C ||f_{1,\sharp}||_{p_1}^{4/3} ||f_{2,\sharp}||_{p_2}^{4/3} \log(||f_{1,\sharp}||_{p_1}^{-1}) \log(||f_{2,\sharp}||_{p_2}^{-1}) = o(||\mathbf{f}||_{\mathbf{p}}^2) \end{aligned}$$

□

Lemma 3.8. *For all $f \in L^{p_1}(\mathbb{R}^{2d})$*

$$\iint f(x) g_2(y) g_3(x+y) \sigma(Ax, Ay) dx dy = 0 \tag{3.25}$$

The conclusion also applies to the same integral with (g_1, f, g_3) or (g_1, g_2, f) in place of (f, g_2, g_3) (with $f \in L^{p_j}$ for the appropriate $j \in \{1, 2, 3\}$).

Proof. Since $\sigma(Ax, Ay) = x^T A^T J A y$ is an antisymmetric bilinear form, we may diagonalize $A^T J A$ as $Q^T \Sigma Q$ for some orthogonal Q and

$$\Sigma = \begin{pmatrix} 0 & a_1 & ... & 0 & 0 \\ -a_1 & 0 & ... & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & ... & 0 & a_d \\ 0 & 0 & ... & -a_d & 0 \end{pmatrix}, \tag{3.26}$$

where $a_k \in \mathbb{R}$ and $\pm a_k i$ are the eigenvalues of $A^T J A$. Since $g_j(x) = e^{-\pi p_j' |x|^2}$, g_2 and g_3 remain unchanged under an orthogonal change of coordinates. Thus, we may write the term in question as

$$\iint f(Qx) g_2(y) g_3(x+y) \sum_{k=1}^{d} a_k (x_{2k-1} y_{2k} - x_{2k} y_{2k-1}) dx dy. \tag{3.27}$$

Since $f(x)$ is an arbitrary function of x, $f(Qx)$ is also an arbitrary function of x, so it suffices to show that

$$\int g_2(y) g_3(x+y) \sum_{k=1}^{d} a_k (x_{2k-1} y_{2k} - x_{2k} y_{2k-1}) dy = 0 \tag{3.28}$$

for all $x \in \mathbb{R}^{2d}$.

By linearity and permutation of coordinates, it suffices to show that

$$\int g_2(y) g_3(x+y)(x_1 y_2 - x_2 y_1) dy = 0. \tag{3.29}$$

Writing $e^{-\pi p_j' |w|^2} = e^{-\pi p_j' (w_1^2 + w_2^2)} e^{-\pi p_j' (w_3^2 + ... + w_{2d}^2)}$, the above integral factors into

$$\int g_2(y_1, y_2) g_3(x_1 + y_1, x_2 + y_2)(x_1 y_2 - x_2 y_1) dy_1 dy_2 \cdot \int g_2(\tilde{y}) g_3(\tilde{x} + \tilde{y}) d\tilde{y}, \tag{3.30}$$

where $x = (x_1, x_2, \tilde{x}), y = (y_1, y_2, \tilde{y})$, and through abuse of notation, $g_j(w) = e^{-p_j' |w|^2}$ for w in any dimension. It now suffices to show the first factor is zero.

Expanding this factor gives

$$\begin{aligned} x_1 \int y_2 g_2(y_2) g_3(x_2 + y_2) dy_2 \cdot \int g_2(y_1) g_3(x_1 + y_1) dy_1 \\ - x_2 \int y_1 g_2(y_1) g_3(x_1 + y_1) dy_1 \cdot \int g_2(y_2) g_3(x_2 + y_2) dy_2. \end{aligned} \tag{3.31}$$

An elementary computation shows that $g_2 * g_3 = Cg_1$ and $\int y g_2(y) g_3(x+y)dy = C' x g_1(x)$, hence the above becomes

$$x_1 \cdot C' x_2 g_1(x_2) \cdot C g_1(x_1) - x_2 \cdot C' x_1 g_1(x_1) \cdot C g_1(x_2) = 0. \tag{3.32}$$

□

If S is a list of parameters, let $A \approx_S B$ mean there exists a $C > 0$ depending only on elements of S such that $A \leq CB$ and $B \leq CA$.

Lemma 3.9. *For* $\mathbf{g}$ *and* A *as above,*

$$\iint g_1(x) g_2(y) g_3(x+y) \sigma^2(Ax, Ay) dxdy \approx_{d,\mathbf{p}} ||A^T J A||^2. \tag{3.33}$$

Proof. As in the proof of Lemma 3.8, one may use an orthogonal change of coordinates to reduce to the computation of

$$\iint g_1(x) g_2(y) g_3(x+y) \left[\sum_{k=1}^{d} a_k (x_{2k-1} y_{2k} - x_{2k} y_{2k-1}) \right]^2 dxdy. \tag{3.34}$$

Expanding the square gives

$$\sum_{j,k=1}^{d} a_j a_k \iint g_1(x) g_2(y) g_3(x+y) (x_{2j-1} y_{2j} - x_{2j} y_{2j-1})(x_{2k-1} y_{2k} - x_{2k} y_{2k-1}) dxdy.$$

By factoring the g_j and computing the above integrals two coordinates at a time as in the proof of Lemma 3.8, one finds that the cross terms are zero. Thus, the original integral is equal to a function to d and $\mathbf{p}$ alone times $\sum_{k=1}^{d} a_k^2$. Recall that $\pm a_k i$ are the eigenvalues of $A^T J A$, so $||A^T J A||^2 = \max_k |a_k|^2$ and the two expressions are equivalent. □

At this point, it is tempting to expand $\mathcal{T}_A(\mathbf{g}+\mathbf{f})$, using the previous four lemmas to treat the $(\mathcal{T}_A - \mathcal{T}_0)$ terms (to get $-c||A^T J A||^2$) and Theorem 3.1 to treat the $\mathcal{T}_0$ terms (and get $A_{\mathbf{p}}^{2d} - c||\mathbf{f}||_{\mathbf{p}}^2$). However, Theorem 3.1 may only be applied directly when the perturbative terms f_j represent the projective distance from the orbit of the original functions to $\mathbf{g}$. The subtle difference here is that the f_j which represent the minimum value of $||\mathbf{f}||_{\mathbf{p}}^2$ may not be the same functions which represent the minimum value of $||\mathbf{f}||^2 + ||A^T J A||^2$.

For this reason, we will delve somewhat into the proof of Theorem 3.1 and show that it is possible to obtain the same circumstances which lead to a $-c||\mathbf{f}||^2$ decay.

3.4 Balancing Lemma

For $t > 0$ and $n = 0, 1, 2, ...,$ let $P_n^{(t)}$ denote the real-valued polynomial of degree n with positive leading coefficient and $||P_n^{(t)} e^{-t\pi x^2}||_{L^2(\mathbb{R})} = 1$ which is orthogonal to $P_k^{(t)} e^{-t\pi x^2}$ for all $0 \le k < n$.

For $d > 1$, $\alpha = (\alpha_1, ..., \alpha_d) \in \{0, 1, 2, ...\}^d$, and $x = (x_1, ..., x_d) \in \mathbb{R}^d$, define

$$P_\alpha^{(t)}(x) = \prod_{k=1}^{d} P_{\alpha_k}^{(t)}(x_k). \tag{3.35}$$

Let $\tau_j = \frac{1}{2} p_j p_j'$ $(j = 1, 2, 3)$. In [17], the following is proved en route to the main theorem.

Proposition 3.10. *Let $\delta_0 > 0$ be sufficiently small. There exists $c, \tilde{c} > 0$ and a choice of $\eta > 0$ in the $f_j = f_{j,\sharp} + f_{j,\flat}$ decomposition such that the following holds. Suppose $||\mathbf{f}||_{\mathbf{p}} < \delta_0$ and f_j satisfy the following orthogonality conditions:*

- $\langle Re(f_j), P_\alpha^{(\tau_j)} g_j^{p_j - 1} \rangle = 0$ *whenever $\alpha = 0$, $|\alpha| = 1$ and $j \in \{1, 2\}$, or $|\alpha| = 2$ and $j = 3$.*
- $\langle Im(f_j), P_\alpha^{(\tau_j)} g_j^{p_j - 1} \rangle = 0$ *whenever $\alpha = 0$ or $|\alpha| = 1$ and $j = 3$.*

Then,

$$\frac{\mathcal{T}_0(\mathbf{g} + \mathbf{f})}{\prod_j ||g_j + f_j||_{p_j}} \le A_{\mathbf{p}}^{2d} - c||\mathbf{f}||_{\mathbf{p}}^2 - \tilde{c} \sum_j ||f_{j,\flat}||_{p_j}^{p_j}. \tag{3.36}$$

The above proposition is not stated as an explicit result of [17]. However, (3.36) is, in effect, the penultimate line of the proof of Theorem 3.1 in Section 8 of [17]. (The one difference is that $c||\mathbf{f}||_{\mathbf{p}}^2$ is replaced by $\sum_j ||f_{j,\sharp} g_j^{(p_j-2)/2}||_2^2$ in the line in [17], though it is shown the latter majorizes a constant multiple of the former.)

We cite this particular intermediate result in order to take advantage of the $f_j = f_{j,\sharp} + f_{j,\flat}$ decomposition. The terms in Lemma 3.7 involve $f_{j,\sharp}$ in place of f_j so (3.36) is used to deal with the case that $f_{j,\flat}$ makes up a significant portion of the L^{p_j} norm of f_j.

The goal of this section is to reduce to the situation in which the hypotheses of Proposition 3.10 apply. This is done through the use of the following balancing lemma.

Lemma 3.11 (Balancing Lemma)**.** *Let $d \ge 1$ and $\mathbf{p} \in (1, 2]^3$ with $\sum_j p_j^{-1} = 2$. There exists $\delta_0 > 0$ such that if $||F_j - g_j||_{p_j} \le \delta_0$, $||A^T J A|| \le \delta_0$, and $\langle F_j - g_j, g_j^{p_j - 1} \rangle = 0$,* *then there exist* $v_j \in \mathbb{R}^{2d}$ *satisfying* $v_1+v_2+v_3=0$, $a_j \in \mathbb{C}$, $\xi \in \mathbb{R}^{2d}$, *and a* $(2d)\times(2d)$ *matrix* ψ *such that*

$$\sum_j (|v_j| + |a_j - 1|) + ||\psi - I_{2d}|| + |\xi| \leq C\left[\left(\sum_j ||f_j - g_j||_{p_j}\right)^2 + ||A^TJA||^2\right] \quad (3.37)$$

and the orthogonality conditions of Proposition 3.10 hold for the functions

$$\tilde{F}_j(x) = a_j F_j(\psi(x) + v_j)e^{ix\cdot\xi + iA^TJA\tilde{v_j}\cdot x}, \quad (3.38)$$

where $\tilde{v_1} = v_2, \tilde{v_2} = v_1$, *and* $\tilde{v_3} = 0$.

Proof. Begin by writing $h_j = g_j - F_j$ and $\tilde{h_j} = g_j - \tilde{F_j}$, where $\tilde{F_j} = a_jF_j(\psi(x) + v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x}$ and ψ, v_j, a_j, ξ are to be determined. Letting $a_j = 1+b_j$ and subbing in $f_j = g_j + h_j$,

$$\begin{aligned}\tilde{h_j}(x) &= a_jF_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x} - g_j(x)\\ &= (1+b_j)(g_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x} + h_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x}) - g_j(x).\end{aligned}$$

Writing $\psi(x) = x + \phi(x)$ and taking the two terms involving g_j from above,

$$\begin{aligned}&g_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x} - g_j(x)\\ &\quad= g_j(x)[g_j^{-1}(x)g_j(x+v_j+\phi(x))e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x} - 1]\\ &\quad= g_j(x)(e^{-\pi p_j'[|x+v_j+\phi(x)|^2-|x|^2]}e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x} - 1)\\ &\quad= g_j(x)x\cdot[-2p_j'(\phi(x)+v_j) + i\xi + iA^TJA\tilde{v_j}] + O((||\phi|| + |\mathbf{v}| + |\xi|)^2),\end{aligned}$$

where $O((||\phi|| + |\mathbf{v}| + |\xi|)^2)$ represents the L^{p_j} norm of the remainder term. Substituting back into the initial expression for $\tilde{h_j}$, one finds

$$\begin{aligned}&\tilde{h_j}(x) = a_jh_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x}\\ &+ g_j(x)x\cdot[-2p_j'(\phi(x)+v_j) + i\xi + iA^TJA\tilde{v_j}] + O((||\phi|| + |\mathbf{v}| + |\mathbf{b}| + |\xi|)^2). \quad (3.39)\end{aligned}$$

In computing $\langle \tilde{h_j}, P_\alpha^{(\tau_j)}g_j^{p_j-1}\rangle$, we begin with the main term from (3.39).

$$\begin{aligned}&\langle a_jh_j(\psi(x)+v_j)e^{ix\cdot\xi+iA^TJA\tilde{v_j}\cdot x}, P_\alpha^{(\tau_j)}g_j^{p_j-1}\rangle\\ &\qquad= \langle h_j(\psi(x)+v_j), P_\alpha^{(\tau_j)}g_j^{p_j-1}\rangle + O((|\mathbf{b}| + |\xi| + |\mathbf{v}|)||h_j||_{p_j})\\ &\qquad= |\det(\psi)|^{-1}\int h_j(y)P_\alpha^{(\tau_j)}(\psi^{-1}(y-v_j)g_j^{p_j-1}(\psi^{-1}(y-v_j)dy\\ &\qquad\qquad+ O((|\mathbf{b}| + |\xi| + |\mathbf{v}|)||h_j||_{p_j})\\ &\qquad= \langle h_j, P_\alpha^{(\tau_j)}g_j^{p_j-1}\rangle + O((|\mathbf{b}| + |\xi| + ||\phi|| + |\mathbf{v}|)||h_j||_{p_j}).\end{aligned}$$

Considering the full expression from (3.39),

$$\begin{aligned}\langle \tilde{h_j}, P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle = \langle h_j, P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle \\ + \langle g_j(x) x\cdot[b_j - 2p_j'(\phi(x)+v_j) - i\xi - iA^T JA\tilde{v_j}], P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle \\ + O((||\phi|| + |\mathbf{v}| + |\mathbf{b}| + |\xi|)^2 + (||\phi|| + |\mathbf{v}| + |\mathbf{b}| + |\xi|)||h_j||_{p_j}). \quad (3.40)\end{aligned}$$

In order to complete the proof via the Implicit Function Theorem, it suffices to show that the map

$$(\mathbf{b}, \mathbf{v}, \xi, \phi) \mapsto \langle g_j(x)[b_j - v_j \cdot x - i(\xi + A^T JA\tilde{v_j})\cdot x - 2p_j' x\cdot\phi(x)], P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle \quad (3.41)$$

with (j, α) ranging over the indices specified in Proposition 3.10 and taking the real or imaginary part as specified is invertible.

Since $\{x\cdot\phi(x) : \phi \text{ is a symmetric real } (2d)\times(2d) \text{ matrix}\}$ is precisely the set of symmetric, real, homogeneous, quadratic polynomials on $\mathbb{R}^{2d}$, the map $\phi \mapsto (\langle x\cdot\phi(x)g_3(x), P_\alpha^{(\tau_3)} g_3^{p_3-1}\rangle : |\alpha| = 2)$ is invertible. These inner products vanish when $|\alpha| = 0, 1$.

The contribution from the mapping $(\mathbf{v}, \xi)$ with the constraint $v_1 + v_2 + v_3 = 0$ to $\langle g_j(x)[v_j\cdot x - i(\xi + A^T JA\tilde{v_j})\cdot x], P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle$ ranging over the indices of Proposition 3.10 and taking the real and imaginary parts is also invertible. These products vanish when $|\alpha| = 0, 2$.

Lastly, the contribution from $\langle g_j(x)b_j, P_\alpha^{(\tau_j)} g_j^{p_j-1}\rangle$ indexed over $j = 1, 2, 3$ is in one-to-one correspondence with $\mathbf{b}$ and these inner products vanish when $|\alpha| = 1, 2$. Thus, the maps described in (3.41) is invertible. □

3.5 Putting it All Together

Proof of Theorem 3.2. Let (h_1, h_2, h_3, B) be a 4-tuple with $h_j \in L^{p_j}$ and B an arbitrary $(2d)\times(2d)$ matrix such that $\text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{h}, B), (\mathbf{g}, 0))$ is sufficiently small. By the Balancing Lemma, there exists an element (F_1, F_2, F_3, A) of the orbit of $(\mathbf{h}, B)$ which satisfies the orthogonality conditions of Proposition 3.10. Let $f_j = F_j - g_j$. Since

$$\text{dist}_{\mathbf{p}}(\mathcal{O}_{TC}(\mathbf{h}, B), (\mathbf{g}, 0))^2 \le \max_j ||f_j||_{p_j}^2 + ||A^T JA||^2, \quad (3.42)$$

it suffices to prove that

$$\frac{\mathcal{T}_A(\mathbf{g}+\mathbf{f})}{\prod_j ||g_j + f_j||_{p_j}} \le A_{\mathbf{p}}^{2d} - c\left[\max_j ||f_j||_{p_j}^2 + ||A^T JA||^2\right]. \quad (3.43)$$

By Proposition 3.10,

$$\frac{\mathcal{T}_0(\mathbf{g}+\mathbf{f})}{\prod_j ||g_j+f_j||_{p_j}} \leq A_{\mathbf{p}}^{2d} - c||\mathbf{f}||_{\mathbf{p}}^2 - \tilde{c}\sum_j ||f_{j,\flat}||_{p_j}^{p_j}. \tag{3.44}$$

Thus, it suffices to show that

$$\frac{(\mathcal{T}_A-\mathcal{T}_0)(\mathbf{g}+\mathbf{f})}{\prod_j ||g_j+f_j||_{p_j}} \leq -c||A^TJA||^2 + O((||\mathbf{f}||_{\mathbf{p}}+||A^TJA||)^3). \tag{3.45}$$

We may ignore the product of norms in the denominator by appropriate modification of the constant c. Expanding $(\mathcal{T}_A-\mathcal{T}_0)(\mathbf{g}+\mathbf{f})$ through the multilinearity of $\mathcal{T}_A-\mathcal{T}_0$, one obtains four types of terms. By Lemma 3.8 and Lemma 3.9,

$$\begin{aligned}(\mathcal{T}_A-\mathcal{T}_0)(g_1,g_2,g_3) &= \iint g_1(x)g_2(y)g_3(x+y)(e^{i\sigma(Ax,Ay)}-1)dxdy \\ &= \iint g_1(x)g_2(y)g_3(x+y)(i\sigma(Ax,Ay)-\frac{1}{2}\sigma(Ax,Ay)^2+O(\sigma(Ax,Ay)^3))dxdy \\ &= -c||A^TJA||^2+O(||A^TJA||^3).\end{aligned}$$

By similar application of Lemma 3.8,

$$\begin{aligned}(\mathcal{T}_A-\mathcal{T}_0)(f_1,g_2,g_3) &= \iint f_1(x)g_2(y)g_3(x+y)(i\sigma(Ax,Ay)+O(\sigma(Ax,Ay)^2))dxdy \\ &\leq 0+||A^TJA||^2\int f_1(x)x^2\left[\int y^2g_2(y)g_3(x+y)dy\right]dx \\ &= O(||f_1||_{p_1}||A^TJA||^2)\end{aligned}$$

and likewise for all other terms involving one f_j and two g_j's.

The $(\mathcal{T}_A-\mathcal{T}_0)(f_1,f_2,f_3)$ term is negligible by Lemma 3.6, so only the terms with two f_j's and one g_j remain. Lemma 3.7 only addresses the situation where the f_j are replaced with $f_{j,\sharp}$. However, Proposition 3.10 provides a $-\tilde{c}\sum_j ||f_{j,\flat}||_{p_j}^{p_j}$ term which may be used here. Expanding further and applying Lemma 3.7 and (3.23) gives

$$\begin{aligned}&|(\mathcal{T}_A-\mathcal{T}_0)(f_1,f_2,g_3)| - \tilde{c}\sum_j ||f_{j,\flat}||_{p_j}^{p_j} \\ &\leq o(||\mathbf{f}||_{\mathbf{p}}^2+||A^TJA||^2)+O(||f_{1,\sharp}||_{p_1}||f_{2,\flat}||_{p_2}+||f_{2,\sharp}||_{p_2}||f_{1,\flat}||_{p_1}) - \tilde{c}\sum_j ||f_{j,\flat}||_{p_j}^{p_j}.\end{aligned}$$

If $\sum_j ||f_{j,\flat}||_{p_j}^{p_j}$ is small relative to $||\mathbf{f}||_{\mathbf{p}}^2$, then the above is negligible, as each $||f_{j,\flat}||_{p_j}$ is small. (Specifically, one may split into cases where $||f_{j,\flat}||_{p_j} \geq ||f_j||_{p_j}^{(4-p_j)/2}$ for at least one j or none of the j.) However, if $\sum_j ||f_{j,\flat}||_{p_j}^{p_j}$ is large relative to $||\mathbf{f}||_{\mathbf{p}}^2$, then the last term dominates (as $p_j < 2$), and the above is still negligible.

This holds for the other terms involving two f_j's and one g_j, thus completing the proof of the main theorem.

□

Chapter 4

A Sharpened Inequality for Convolution on the Sphere

4.1 Introduction

Recall from Chapter 1 that a version of convolution on the sphere S^d is given by the trilinear form

$$\mathcal{T}(f,g,h) := \iint_{S^d\times S^d} f(x)g(y)h(x\cdot y)d\sigma(x)d\sigma(y), \tag{4.1}$$

where $f,g : S^d \to \mathbb{R}$, $h : [-1,1] \to \mathbb{R}$ and σ is surface measure on S^d (normalized so $\sigma(S^d) = 1$).

Keeping with the theme of this thesis, we ask for which functions f, g, and h is $\mathcal{T}(f,g,h)$ relatively large? A useful answer involves the *nondecreasing symmetric rearrangement* f^* of a function $f : S^d \to \mathbb{R}$. Writing the coordinates of S^d as $(x_1,...,x_{d+1})$, $f^* : S^d \to \mathbb{R}$ is defined as the unique function (up to sets of measure zero) which depends only on x_{d+1}, is nondecreasing in x_{d+1}, and has the same distribution function as f. (That is, $\sigma(\{f^* > \lambda\}) = \sigma(\{f > \lambda\})$ for all $\lambda \in \mathbb{R}$.)

A classical result of Baernstein and Taylor [3] says the following:

Theorem 4.1. *Let $h : [-1,1] \to [0,\infty)$ be a nondecreasing, bounded, measurable function and let $f,g \in L^1(S^d)$. Then,*

$$\mathcal{T}(f,g,h) \le \mathcal{T}(f^*,g^*,h). \tag{4.2}$$

This is analogous to the Riesz-Sobolev inequality

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} f(x)g(y)h(x+y)dxdy \le \iint_{\mathbb{R}^d\times\mathbb{R}^d} f^*(x)g^*(y)h^*(x+y)dxdy \tag{4.3}$$

on $\mathbb{R}^d$ (where the symmetric decreasing rearrangement is defined as in Chapter 1), except the latter makes no extra assumption on the function h. Further work of Baernstein established the conclusion of Theorem 4.1 without the nondecreasing hypothesis on h in the $d = 1$ case [2]. To the best of the author's knowledge, this remains an open problem in dimensions $d \geq 2$.

In this chapter, we prove a sharpened version of (4.2) in the special case where f, g and h are indicator functions, writing $\mathcal{T}(E_1, E_2, I) = \mathcal{T}(\mathbf{1}_{E_1}, \mathbf{1}_{E_2}, \mathbf{1}_I)$. Here, $E_j \subset S^d$ and $I = [a, 1]$ for some $-1 < a < 1$. Often, $\mathcal{T}(E_1, E_2, I)$ will be written as $\mathcal{T}(E_1, E_2)$ when I is clear from context. E^* will be used to denote the set such that $(\mathbf{1}_E)^* = \mathbf{1}_{E^*}$, or alternatively, the spherical cap of measure $\sigma(E)$ with center $N = (0, ..., 0, 1)$.

Two obstacles arise in constructing the proper statement of our main theorem.

First, observe that $\mathcal{T}(Q(E_1), Q(E_2), I) = \mathcal{T}(E_1, E_2, I)$ for any orthogonal transformation Q on $\mathbb{R}^{d+1}$. In fact, the symmetry group of the operator $\mathcal{T}$ is the orthogonal group $O(d + 1)$. Thus, any sharpened version of (4.2) must account for the case in which E_1, E_2 are spherical caps centered at some point other than N.

For this purpose, define the distance of a pair of sets $\mathbf{E} = (E_1, E_2)$ from the orbit of maximizers to be

$$\operatorname{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) := \inf_{Q \in O(d+1)} \max_{j \in \{1,2\}} |E_j \Delta Q(E_j^*)|, \tag{4.4}$$

where $A \Delta B$ denotes the symmetric difference between the sets A and B and $|A|$ refers to the surface measure of A.

Second, consider for example the case where $I = [0, 1]$ and E_1, E_2 are small spherical caps centered at N. A small perturbation of E_1 and E_2 preserves the property that $x \cdot y > 0$ for all $x \in E_1$ and $y \in E_2$. Therefore, $\mathcal{T}(\mathbf{E})$ remains constant, despite taking the sets further from their rearrangements. A similar problem arises if E_2^* is much bigger than E_1^* and I is a small interval.

To address this issue, let r_i $(i = 1, 2)$ be the spherical radius of the cap E_i^*, that is, the Riemannian distance on S^d from the center of E_i^* to its boundary. If I is the interval $[a, 1]$, then let r_3 be the Riemannian distance between any two points on S^d whose dot product is a. We say that $(\sigma(E_1), \sigma(E_2), I)$ (or (E_1, E_2, I)) is *strictly admissible* if $r_i + r_j > r_k$ for all permutations (i, j, k) of $(1, 2, 3)$.

Our main theorem is the following.

Theorem 4.2. *Let $d \geq 1$ and K be a compact subset of the set of strictly admissible triples (e_1, e_2, a) with $0 < e_1, e_2 < 1, -1 < a < 1$. Then, there exists $c > 0$ such that for all $(e_1, e_2, a) \in K$ and pairs of Lebesgue measurable subsets $E_j \subset S^d$ with $\sigma(E_j) = e_j$,*

$$\mathcal{T}(E_1, E_2, I) \leq \mathcal{T}(E_1^*, E_2^*, I) - c\,\mathit{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))^2, \tag{4.5}$$

where $I = [a, 1]$.

Results of this type were proved for $\mathbb{R}$ in [13] and for $\mathbb{R}^d$ in [15].

The $d = 1$ case of Theorem 4.2 is established by Christ and Iliopoulou in the more general case where I is replaced with an arbitrary subset of $[-1, 1]$ [18].

While Theorem 4.2 will be proven focusing on a single arbitrary triple (e_1, e_2, a), one may check that uniformity holds at each step.

To the best of the author's knowledge, the following corollary is new for dimensions $d \geq 2$.

Corollary 4.3. *Let* $d \geq 1$. *Fix a strictly admissible triple* (e_1, e_2, a) *and let* $I = [a, 1]$. *Up to difference of Lebesgue null sets, the maximizers of* $\mathcal{T}(E_1, E_2, I)$ *among* $E_j \subset S^d$ *with* $\sigma(E_j) = e_j$ *are precisely* $\{(Q(E_1^*), Q(E_2^*)) : Q \in O(d+1)\}$.

Again, the $d = 1$ case is established in [18]. A result of this type for $\mathbb{R}^d$ was proven in [10].

The proof of Theorem 4.2 largely follows the proof of the main theorem of [15]. Similar techniques were used in work by Bianchi and Egnell [7]. (See also [19].) The strategy is to expand about (E_1^*, E_2^*) and analyze the quadratic term, reducing to functions supported on the boundaries of E_1^* and E_2^*. From here, the quadratic form may be diagonalized by spherical harmonics and a balancing lemma is used to eliminate those of degree 1.

However, two distinct challenges arise in following this method. First, in this method one reduces to the case of small perturbations, that is, when $\text{dist}((\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) \leq \delta_0$. In [15] and [19], this is done through the use of a continuous flow which takes arbitrary sets to maximizers and under which the functional is nondecreasing. The flow is stopped at precisely the time which the distance from the maximizers is δ_0. In the case of spherical convolution, no such flow is known to exist in dimensions $d \geq 2$. (See [18] for a flow in $d = 1$.)

As an alternative, we use the reflection method deployed in [3] which is used to prove (4.2) by transforming the initial sets under a sequence of reflections about hyperplanes. While a general sequence of reflections may not give rise to sets at a distance δ_0 from their rearrangements, one may modify the sequence to produce the desired distance.

Second, it is not inherently clear how to complete the proof once one reduces to the case of particular sets E_j determined by spherical harmonics. To address this issue in [15], Christ uses the one-dimensional sharpened Riesz-Sobolev [13] and Steiner symmetrization. Given the geometry of the sphere, this part of the proof is completed by induction, applying the sharpened inequality on S^{d-1} to horizontal slices of S^d.

4.2 Reduction to Small Perturbations

In this section, we show that it suffices to prove Theorem 4.2 in the case where $\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))$ is small. Specifically, Theorem 4.2 will follow from Proposition 4.4.

Proposition 4.4. *Let $d \geq 1$ and (e_1, e_2, a) be a strictly admissible triple. Let $I = [a, 1]$. Then, there exist $\delta_0, c > 0$ such that for all Lebesgue measurable sets $E_j \subset S^d$ with $\sigma(E_j) = e_j$ $(j = 1, 2)$,*

$$\mathcal{T}(E_1, E_2, I) \leq \mathcal{T}(E_1^*, E_2^*, I) - c\,dist(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))^2 \tag{4.6}$$

whenever $dist(\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) \leq \delta_0$.

Given an oriented hyperplane H in $\mathbb{R}^{d+1}$ which passes through the origin, let H^+ denote the half-space determined by H and the positive orientation of H, and let H^- denote the complement of H^+. Let ρ_H denote reflection across H.

Construct the set $E_H \subset S^d$ as follows. If $x \in H^+$, then $x \in E_H$ if either $x \in H^+ \cap E$ or $\rho_H(x) \in H^- \cap E$. If $x \in H^-$ then $x \in E_H$ if $x \in H^- \cap E$ and $\rho_H(x) \in H^+ \cap E$.

A useful formula for E_H is

$$E_H = [H^+ \cap (E \cup \rho_H(E)] \cup [H^- \cap E \cap \rho_H(E)]. \tag{4.7}$$

Given oriented hyperplanes H_j, the notation $E_{H_1 \cdots H_n}$ will be used to denote $(\cdots(E_{H_1})_{H_2}\cdots)_{H_n}$.

The following two lemmas were proven in [3] in more general context. (While they were stated for continuous functions in place of sets, a standard approximation argument recovers the conclusion for sets. Lemma 4.6 is implicit in the proof of Theorem 4.1 found in [3]. Furthermore, we allow for a larger class of reflections, though their results extend trivially.)

Lemma 4.5. *Let $E, F \subset S^d$ and H be an oriented hyperplane in $\mathbb{R}^{d+1}$ through the origin. Let E_H and F_H be defined as above. Then,*

1. $|E_H| = |E|$.
2. $\mathcal{T}(E, F) \leq \mathcal{T}(E_H, G_H)$.

Lemma 4.6. *Given sets $F, G \subset S^d$, there exists a sequence of oriented hyperplanes $H_1, ..., H_n, ...$ through the origin such that $|F_{H_1 \cdots H_n} \Delta F^*| \to 0$ and $|G_{H_1 \cdots H_n} \Delta G^*| \to 0$ as $n \to \infty$.*

Proof of Prop. 4.4 $\Rightarrow$ Thm. 4.2. If $\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) \leq \delta_0$, we are done, so assume that $\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) > \delta_0$.

Let $H_1, ..., H_n, ...$ be a sequence of oriented hyperplanes guaranteed by Lemma 4.6 and let $\mathbf{E}_n = (E_{1_{H_1 \cdots H_n}}, E_{2_{H_1 \cdots H_n}})$.

Suppose there exists an n_0 such that $\mathrm{dist}(\mathbf{E}_{n_0}, \mathcal{O}(\mathbf{E}^*)) \in (\delta_0/100, \delta_0)$. Then, by Lemma 4.6, Proposition 4.4, and the finite measure of S^d,

$$\mathcal{T}(\mathbf{E}) \leq \mathcal{T}(\mathbf{E}_{n_0}) \leq \mathcal{T}(\mathbf{E}^*) - c\delta_0^2/10^4 \leq \mathcal{T}(\mathbf{E}^*) - \tilde{c}\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))^2 \tag{4.8}$$

for some $\tilde{c} > 0$ by the boundedness of $\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))$. (The distance is at most $\sigma(S^d)$.)

Now suppose the contrary, that there exists n_0 such that $\mathrm{dist}(\mathbf{E}_{n_0 - 1}, \mathcal{O}(\mathbf{E}^*)) > \delta_0$, but $\mathrm{dist}(\mathbf{E}_{n_0}, \mathcal{O}(\mathbf{E}^*)) < \delta_0/100$. Our goal is to modify H_{n_0} so that $\mathrm{dist}(\mathbf{E}_{n_0}, \mathcal{O}(\mathbf{E}^*))$ lies in the interval $(\delta_0/100, \delta_0)$.

Denote $\mathbf{F} = \mathbf{E}_{n_0 - 1}$ and $H_0 = H_{n_0}$ so $\mathrm{dist}(\mathbf{F}_H, \mathcal{O}(\mathbf{F}^*)) < \delta_0/100$.

Suppose for the moment that there exists an oriented hyperplane $\tilde{H}$ such that $\mathrm{dist}(\mathbf{F}_{\tilde{H}}, \mathcal{O}(\mathbf{F}^*)) > \delta_0/100$. Let $H(t)$ $(0 \leq t \leq 1)$ be a continuous path in the space of oriented hyperplanes through the origin such that $H(0) = H_0$ and $H(1) = \tilde{H}$. If $\mathrm{dist}(\mathbf{F}_{H(t)}, \mathcal{O}(\mathbf{F}^*))$ is continuous in t, one may apply the Intermediate Value Theorem to obtain t_0 such that $\mathrm{dist}(\mathbf{F}_{H(t_0)}, \mathcal{O}(\mathbf{F}^*)) \in (\delta_0/100, \delta_0)$. Thus, replacing H_{n_0} with $H(t_0)$ reduces to the previously addressed case.

To establish continuity of $\mathrm{dist}(\mathbf{F}_{H(t)}, \mathcal{O}(\mathbf{F}^*))$ in t, note that by (4.7), $|A_H \Delta Q(A^*)|$ is jointly continuous in H and Q for any set A. (To prove this, one may use outer regularity of surface measure to reduce to the case where A is a finite union of balls, for which the statement is obvious.) It follows that

$$\mathrm{dist}(\mathbf{F}_H, \mathcal{O}(\mathbf{F}^*)) = \inf_{Q \in O(d)} \max_{j=1,2} |(F_j)_H \Delta Q(F_j^*)| \tag{4.9}$$

is continuous in H on the image of $H(t)$.

It remains to establish the existence of such an $\tilde{H}$. Let $F = F_{j_0}$, where $\max_{j=1,2}$ is obtained with $j = j_0$; that is, $\mathrm{dist}(\mathbf{F}, \mathcal{O}(\mathbf{F}^*)) = \inf_Q |F_{j_0} \Delta Q(F_{j_0})|$. We begin with a series of reductions. By applying an orthogonal transformation on F, one may suppose without loss of generality that H_0 is the hyperplane determined by $x_d = 0$ oriented toward the positive x_d-axis. Thus, choosing Q such that $|F_{H_0} \Delta Q(F^*)| < \delta_0/100$, the center of $Q(F^*)$ is contained in $H^+ = \{x_d \geq 0\}$. Taking complements, it suffices to prove the claim for $|F| \leq 1/2$. By possibly replacing H_0 with its opposite orientation, suppose $|F \cap H_0^+| \geq 1/2|F|$.

Since $\mathrm{dist}(F, \mathcal{O}(F^*)) > \delta_0$ and $|F_{H_0} \Delta Q(F^*)| < \delta_0/100$, we have $|F \Delta F_{H_0}| > 99\delta_0/100$, hence there exists a set G such that $|G| > \delta_0/25$, $\rho_{H_0}(G) \subset Q(F^*)$, $G \cap$

$Q(F^*) = \emptyset$, and a small rotation of $\rho_{H_0}(G)$ intersects F nontrivially. Hence, we may choose $\tilde{H}$ to be a rotation of H_0 of angle $O(\delta_0)$ in the $x_d x_{d+1}$-plane such that $|G \cap F_{\tilde{H}}| > \delta_0/100$.

Let G' be the subset of $Q(F^*)$ of size $\delta_0/25$ which is furthest from G. If $|G' \cap F| > \delta_0/50$, then we are done since $G \cap Q(E^*) = \emptyset$. Else, since $|Q(F^*) \cap F_{\tilde{H}}| < \delta_0/100$, $|\rho_{\tilde{H}}(G') \cap F| > \delta_0/50$ and we are done because no orthogonal transformation of F^* may contain $\rho_{\tilde{H}}(G') \cap F, G$, and $\rho_{\tilde{H}}(G)$.

□

4.3 Reductions to Near and on the Boundary

At this point, we have reduced to the case of small perturbations. Hence, we assume that $\text{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*)) \le \delta_0$ and that (E_1, E_2, I) is strictly admissible.

Reduction to Near the Boundary

Replace the coordinates $(x_1, ..., x_{d+1})$ of S^d with (θ, t), where $t = x_{d+1}$ and θ represents the spherical coordinates on the horizontal S^{d-1} slices of S^d. For $j = 1, 2$, let $B_j = E_j^*$ and let h_j be the value of t_j for the points of ∂B_j. A useful fact is that $d\sigma(x) = (1 - t^2)^{\frac{d}{2}} dt d\mu(\theta)$, where μ is the surface measure on S^{d-1}, normalized so $\mu(S^{d-1}) = 1$.

Choose $Q \in O(d + 1)$ such that $\max_j |Q(E_j)\Delta E_j^*| = \text{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))$ and replace $\mathbf{E}$ with $Q(\mathbf{E})$. (Note that equality *is* attained in the definition of distance since it is the minimum of a continuous function on a compact set.)

Also, let $f_j = \mathbf{1}_{E_j} - \mathbf{1}_{B_j}$.

Expand

$$\mathcal{T}(\mathbf{E}) = \mathcal{T}(\mathbf{1}_{B_1} + f_1, \mathbf{1}_{B_2} + f_2) = \mathcal{T}(\mathbf{E}^*) + \sum_{j=1}^{2} \langle K_j, f_j \rangle + \mathcal{T}(f_1, f_2), \tag{4.10}$$

where $\langle \cdot, \cdot \rangle$ is the L^2 inner product on S^d and

$$K_j(x) = \int_{S^d} \mathbf{1}_{B_k}(y) \mathbf{1}_I(x \cdot y) d\sigma(y) \tag{4.11}$$

with the notation $\{j, k\} = \{1, 2\}$. Each K_j is nonnegative and symmetric (depending only on t). Writing $K_j(x) = K_j(t)$, K_j is increasing in t.

The strict admissibility hypothesis is equivalent to the statement that $K_j'(h_j) > 0$ for $j = 1, 2$. A related equivalent assertion is that for each $x \in \partial B_j$, there exist $y, y' \in B_k$ $(k \ne j)$ such that $x \cdot y > a$ and $x \cdot y' < a$.

Since $\int f_j = 0$,

$$\begin{aligned}\langle K_j, f_j\rangle &= \int (K_j(x) - K_j(h_j))f_j(x)d\sigma(x)\\ &= -\int |K_j(x) - K_j(h_j)| \cdot |f_j(x)|d\sigma(x),\end{aligned}$$

as $K_j(x) - K_j(h_j)$ and $-f_j(x)$ are both nonnegative on B_j and nonpositive on B_j^c.

Let λ be a large positive constant independent of δ to be chosen later. For sufficiently small δ (say $\leq \delta_0$), $K_j'(h_j) > 0$ implies

$$\begin{aligned}\langle K_j, f_j\rangle &\leq -c\lambda\delta \int_{|t-h_j|\geq\lambda\delta} |f_j(x)|d\sigma(x)\\ &= -c\lambda\delta|\{(\theta, t) \in E_j\Delta B_j : |t - h_j| \geq \lambda\delta\}|.\end{aligned}$$

The above term is linear in δ, while $\mathcal{T}(f_1, f_2)$ is quadratic in δ. For this reason, we reduce to the case in which $E_j\Delta B_j \subset \{(\theta, t) : |t - h_j| \leq \lambda\delta\}$, with the formal argument appearing below.

By an argument found in [15], for each $j = 1, 2$, there exists a set $E_j^\dagger$ such that

1. $|E_j^\dagger| = |E_j|$.
2. $E_j\Delta B_j$ is the disjoint union of $E_j^\dagger\Delta B_j$ and $E_j\Delta E_j^\dagger$.
3. $\{(\theta, t) \in E_j\Delta B_j : |t - h_j| > \lambda\delta\} \subset E_j^\dagger\Delta E_j$.
4. $|E_j^\dagger\Delta E_j| \leq 2|\{(\theta, t) \in E_j\Delta B_j : |t - h_j| > \lambda\delta\}|$.

Lemma 4.7. *Let $d \geq 1$ and let (e_1, e_2, a) be a strictly admissible triple and let $I = [a, 1]$. Then, there exist $\lambda < \infty$ and $\delta_0, c > 0$ with the following property. If $E_j \subset S^d$ are Lebesgue measurable sets such that $\sigma(E_j) = e_j$ for $j \in \{1, 2\}$, $\max_{j=1,2} |E_j\Delta E_j^*| \leq \delta_0$, and $\mathbf{E}^\dagger$ is defined as above, then*

$$\mathcal{T}(\mathbf{E}, I) \leq \mathcal{T}(\mathbf{E}^*, I) - c\lambda \sum_{i=1}^{2} |E_i\Delta E_i^*| \cdot \sum_{j=1}^{2} |E_j\Delta E_j^\dagger|. \tag{4.12}$$

Proof. Let $\delta = \max_{j=1,2} |E_j\Delta E_j^*| \leq \delta_0$. Let $f_j^\dagger = \mathbf{1}_{E_j^\dagger} - \mathbf{1}_{B_j}$ and write

$$\mathbf{1}_{E_j} = \mathbf{1}_{B_j} + f_j^\dagger + \tilde{f}_j, \tag{4.13}$$

where $\tilde{f}_j = \mathbf{1}_{E_j} - \mathbf{1}_{E_j^\dagger}$.

Expanding $\mathcal{T}(\mathbf{1}_{B_1} + f_1^\dagger + \tilde{f}_1, \mathbf{1}_{B_2} + f_2^\dagger + \tilde{f}_2)$, one obtains 9 terms. The 4 terms which do not contain an $\tilde{f}_j$ recombine to form $\mathcal{T}(\mathbf{E}^\dagger)$.

There are 2 terms of the form $\langle K_j, \tilde{f}_j \rangle$. By property (1) of $E_j^\dagger$, $\int f_j = 0$. Thus, by the previous discussion and property (3) of $E_j^\dagger$, the sum of these two terms is less than or equal to $-c\lambda\delta \sum_j |E_j \Delta E_j^\dagger|$.

The remaining 3 terms each contain two or more $\tilde{f}_j$ or an $\tilde{f}_j$ and an $f_k^\dagger$. By the inequality

$$\mathcal{T}(E_1, E_2) \leq |E_1| \cdot |E_2| \tag{4.14}$$

and properties (2) and (4) of $E_j^\dagger$, each of these terms is

$$O(\max_j |E_j \Delta E_j^*| \cdot \max_k |E_k^\dagger \Delta E_k|) = O(\delta \max_k |E_k^\dagger \Delta E_k|). \tag{4.15}$$

Putting this together and taking λ large enough, we obtain

$$\begin{aligned} \mathcal{T}(\mathbf{E}) &\leq \mathcal{T}(\mathbf{E}^\dagger) - c\lambda\delta \sum_j |E_j \Delta E_j^\dagger| + O(\delta \max_k |E_k^\dagger \Delta E_k|) \\ &\leq c\lambda \sum_{i=1}^{2} |E_i \Delta E_i^*| \cdot \sum_{j=1}^{2} |E_j \Delta E_j^\dagger|. \end{aligned}$$

The claim follows from the conclusion of Theorem 4.1, that $\mathcal{T}(\mathbf{E}^\dagger) \leq \mathcal{T}(\mathbf{E}^*)$.

□

If $\max_j |E_j \Delta E_j^\dagger| \geq \frac{1}{2} \max_j |E_j \Delta E_j^*|$, then the conclusion of Theorem 4.2 follows immediately from Lemma 4.7.

If $\max_j |E_j \Delta E_j^\dagger| \leq \frac{1}{2} \max_j |E_j \Delta E_j^*|$, Lemma 4.7 still gives $\mathcal{T}(\mathbf{E}) \leq \mathcal{T}(\mathbf{E}^\dagger)$, so it suffices to prove

$$\mathcal{T}(\mathbf{E}^\dagger) \leq \mathcal{T}(\mathbf{E}^*) - c\mathrm{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^\dagger))^2. \tag{4.16}$$

Reduction to the Boundary

Letting $f_j = \mathbf{1}_{E_j} - \mathbf{1}_{B_j}$ as before, define the functions $f_j^\pm$ with values in $\{0,1\}$ by $f_j^+ = \mathbf{1}_{E_j \setminus B_j}$ and $f_j^- = \mathbf{1}_{B_j \setminus E_j}$. Note that $f_j = f_j^+ - f_j^-$.

Define $F_j^\pm \in L^2(S^{d-1}) = L^2(S^{d-1}, \mu)$ by

$$F_j^\pm(\theta) := \int_{-1}^{1} f_j^\pm(\theta, t)(1-t^2)^{\frac{d}{2}} dt. \tag{4.17}$$

We say F_j is the function *associated* to the set E_j.

By the reduction of the previous subsection, $E_j \Delta B_j \subset \{x : |t - h_j| \leq \lambda\delta\}$. Thus,

$$|E_j \Delta B_j|^2 \approx ||F_j^+||^2_{L^2(S^{d-1})} + ||F_j^-||^2_{L^2(S^{d-1})} \tag{4.18}$$

and it suffices to establish a bound of the form

$$\mathcal{T}(\mathbf{E}) \leq \mathcal{T}(\mathbf{E}^*) - c \sum_{j=1}^{2} \left(||F_j^+||^2_{L^2(S^{d-1})} + ||F_j^-||^2_{L^2(S^{d-1})} \right). \tag{4.19}$$

Define the quadratic form $\mathcal{Q}$ on $L^2(S^{d-1})$ by

$$\mathcal{Q}(F, G) := \iint_{(S^{d-1})^2} F(\theta_1) G(\theta_2) \mathbf{1}_I \left(\sqrt{1 - h_1^2}\sqrt{1 - h_2^2}\theta_1 \cdot \theta_2 + h_1 h_2 \right) d\mu(\theta_1) d\mu(\theta_2).$$

Let $\gamma_j = K_j'(h_j)$, which is positive by the strict admissibility hypothesis.

Proposition 4.8. *Under the hypotheses from our reductions,*

$$\mathcal{T}(\mathbf{E}) \leq \mathcal{T}(\mathbf{E}^*) - \frac{1}{2} \sum_{j=1}^{2} \gamma_j (1 - h_j^2)^{-d/2} (||F_j^+||^2_{L^2} + ||F_j^-||^2_{L^2}) + \mathcal{Q}(F_1, F_2) + O(\delta^3). \tag{4.20}$$

Expand $\mathcal{T}(\mathbf{1}_{B_1} + f_1, \mathbf{1}_{B_2} + f_2)$, obtaining four terms. Proposition 4.8 is the immediate result of the following two lemmas.

Lemma 4.9. *For $j = 1, 2$,*

$$\langle K_j, f_j \rangle \leq -\frac{1}{2} \gamma_j (1 - h_j^2)^{-d/2} (||F_j^+||^2_{L^2} + ||F_j^-||^2_{L^2}) + O(\delta^3). \tag{4.21}$$

Proof. Since $K_j(t)$ is twice continuously differentiable in a neighborhood of $t = h_j$, we may write

$$\langle K_j, f_j \rangle = \int_{S^{d-1}} \int_{-1}^{1} (K_j(h_j) + \gamma_j (t - h_j) + O(\delta^2)) f_j(\theta, t) (1 - t^2)^{d/2} dt d\sigma(\theta). \tag{4.22}$$

Expanding the above gives three integrals. First,

$$\int_{S^{d-1}} \int_{-1}^{1} K_j(h_j) f_j(\theta, t) (1 - t^2)^{d/2} dt d\mu(\theta) = 0 \tag{4.23}$$

since $\int f_j = 0$. Next,

$$\int_{S^{d-1}} \int_{-1}^{1} O(\delta^2) f_j(\theta,t)(1-t^2)^{d/2} dt d\sigma(\theta) = O(\delta^2)||f_j||_{L^1} = O(\delta^3) \tag{4.24}$$

Lastly, factor out the γ_j, split $f_j = f_j^+ - f_j^-$, and consider the integral

$$\int_{S^{d-1}} \int_{-1}^{1} (t-h_j) f_j^+(\theta,t)(1-t^2)^{d/2} dt d\mu(\theta). \tag{4.25}$$

For each θ, the support of $f_j^+(\theta,t)$ is contained in $\{t \le h_k\}$. Among functions g_j such that $\int_{-1}^1 g_j(\theta,t)(1-t^2)^{d/2}dt = F_j^+(\theta)$ and supp $g \subset \{t \le h_j\}$,

$$\int_{S^{d-1}} \int_{-1}^{1} (t-h_j) g(\theta,t)(1-t^2)^{d/2} dt d\mu(\theta) \tag{4.26}$$

is maximized when $g_j(\theta,t) = \mathbf{1}_{[h_j - h(\theta), h_j]}(t)$. Here, the function $h(\theta)$ is defined recursively by

$$F_j^+(\theta) = \int_{h_j - h(\theta)}^{h_j} (1-t^2)^{d/2} dt. \tag{4.27}$$

Since $h(\theta) = O(\delta)$,

$$h(\theta) = (1-h_j^2)^{-d/2} F_j^+(\theta) + O(\delta F_k^+(\theta)). \tag{4.28}$$

Thus, using the fact that $||F_j^+||_{L^\infty} = O(\delta)$,

$$\begin{aligned}
\int_{-1}^{1} (t-h_j) f_j^+(\theta,t)(1-t^2)^{d/2} dt &\le \int_{h_j-h(\theta)}^{h_j} (t-h_j)(1-t^2)^{d/2} dt \\
&= -\frac{1}{2} h(\theta)^2 (1-h_j^2)^{d/2} + O(h(\theta)^3) \\
&= -\frac{1}{2}(1-h_j^2)^{-d/2} F_j^+(\theta)^2.
\end{aligned}$$

Plugging back into the original integral from (4.23), we find that

$$\gamma_j \int_{S^{d-1}} \int_{-1}^{1} (t-h_j) f_j^+(\theta,t)(1-t^2)^{d/2} dt d\mu(\theta) \le -\frac{1}{2}\gamma_j (1-h_j^2)^{-d/2} ||F_j^+||_{L^2}^2 + O(\delta^3).$$

A similar result holds for $F_j^-(\theta)$ and f_j. □

Lemma 4.10. *$\mathcal{T}(f_1, f_2) = \mathcal{Q}(F_1, F_2) + O(\delta^3)$.*

Proof. Rewriting in (θ, t)-coordinates, the left hand side is equal to

$$\int_{(S^{d-1})^2} \iint_{[-1,1]^2} f_1(\theta_1, t_1) f_2(\theta_2, t_2) \mathbf{1}_I \left(\sqrt{1-t_1^2}\sqrt{1-t_2^2}\theta_1 \cdot \theta_2 + t_1 t_2 \right) \times (1-t_1^2)^{d/2}(1-t_2^2)^{d/2} dt_1 dt_2 d\mu(\theta_1) d\mu(\theta_2). \quad (4.29)$$

By the definition of F_i, the right hand side is

$$\iint_{(S^{d-1})^2} F_1(\theta_1) F_2(\theta_2) \mathbf{1}_I \left(\sqrt{1-h_1^2}\sqrt{1-h_2^2}\theta_1 \cdot \theta_2 + h_1 h_2 \right) d\mu(\theta_1) d\mu(\theta_2) + O(\delta^3).$$

To compare, it suffices to observe that, since f_i is supported in a $\lambda\delta$-neighborhood of $\{t_i = h_i\}$,

$$\iint_{[-1,1]^2} f_1(\theta_1, t_1) f_2(\theta_2, t_2) \mathbf{1}_I \left(\sqrt{1-t_1^2}\sqrt{1-t_2^2}\theta_1 \cdot \theta_2 + t_1 t_2 \right) (1-t_1^2)^{d/2}(1-t_2^2)^{d/2} dt_1 dt_2 = F_1(\theta_1) F_2(\theta_2) \mathbf{1}_I \left(\sqrt{1-h_1^2}\sqrt{1-h_2^2}\theta_1 \cdot \theta_2 + h_1 h_2 \right)$$

unless $|\sqrt{1-h_1^2}\sqrt{1-h_2^2}\theta_1 \cdot \theta_2 + h_1 h_2 - a| \le C\delta$. However, by the strict admissibility hypothesis (and taking δ sufficiently small), the $\sigma \times \sigma$ measure of the set of pairs (θ_1, θ_2) satisfying this inequality is $O(\delta)$. Since $F_i = O(\delta)$, the contribution of this set to the integral is $O(\delta^3)$. □

4.4 Diagonalization and Balancing Lemma

By Proposition 4.8, it would suffice to prove Theorem 4.2 by showing that

$$\mathcal{Q}(F_1, F_2) \le A \sum_{j=1}^{2} \gamma_j (1-h_j^2)^{d/2} ||F_j||_{L^2}^2 \quad (4.30)$$

for all $F_j \in L^2(S^{d-1})$ satisfying $\int F_j d\sigma = 0$ with constant $A < \frac{1}{2}$. This is because

$$||F_j||_{L^2}^2 = \langle F_j^+ - F_j^-, F_j^+ - F_j^- \rangle = ||F_j^+||_{L^2}^2 + ||F_j^-||_{L^2}^2 - 2\langle F_j^+, F_j^- \rangle \le ||F_j^+||_{L^2}^2 + ||F_j^-||_{L^2}^2.$$

However, (4.30) does not hold for all $F_j \in L^2(S^{d-1})$ satisfying $\int F_j d\sigma = 0$ and $A < \frac{1}{2}$. If it did, then this inequality combined with the above machinery would imply that for all strictly admissible triples (E_1, E_2, I) satisfying $\max_j |E_j \Delta E_j^*| \le \delta_0$ and $E_j \Delta E_j^* \subset \{(\theta, t) : |t - h_j| \le C \max_j |E_j \Delta E_j^*|\}$, that $\mathcal{T}(\mathbf{E}) \le \mathcal{T}(\mathbf{E}^*) - c \max_j |E_j \Delta E_j^*|^2$. However, this conclusion is false for $\mathbf{E} = Q(\mathbf{E}^*)$, where $Q \in O(d+1)$, and $Q(\mathbf{E}^*)$ satisfies all these hypotheses when Q is small. To fix this problem, we will need to use the full strength of

$$\max_j |E_j \Delta E_j^*| = O(\operatorname{dist}(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))). \tag{4.31}$$

We now diagonalize $\mathcal{Q}$ over the spherical harmonics. Lemma 4.11 will use (4.31) to obtain an orthogonality condition on F_j under which (4.30) does hold for $A < 1/2$.

To begin, let $\mathcal{H}_n \subset L^2(S^{d-1})$ denote the space of all spherical harmonics of degree n. Since $\mathcal{Q}$ is a symmetric quadratic form which commutes with rotations, it is diagonalizable over spherical harmonics in the following sense. Let π_n denote the projection of $L^2(S^{d-1})$ onto $\mathcal{H}_n$. Then there exists a compact, self-adjoint operator T on $L^2(S^{d-1})$ such that $\mathcal{Q}(F, G) \equiv \langle T(F), G \rangle$, $T : \mathcal{H}_n \to \mathcal{H}_n$ for all n, and T agrees with a scalar multiple $\lambda = \lambda(n, r_1, r_2, r_3)$ of the identity on $\mathcal{H}_n$.

Since $\int F_j = 0$ for each j, we have $\pi_0(F_j) = 0$ and

$$\mathcal{Q}(F_1, F_2) = \sum_{n=1}^{\infty} Q(\pi_n(F_1), \pi_n(F_2)). \tag{4.32}$$

Let $\mathcal{Q} \circ \pi_n(F_1, F_2) = \mathcal{Q}(\pi_n(F_1), \pi_n(F_2))$. By the compactness of the operator T given above, it suffices to bound

$$|\mathcal{Q} \circ \pi_n(F_1, F_2)| < \frac{1}{2} \sum_{j=1}^{2} \gamma_j (1 - h_j^2)^{d/2} ||\pi_n(F_j)||_{L^2}^2 \tag{4.33}$$

for *each* $n \ge 1$, as the operator norms of $\mathcal{Q} \circ \pi_n$ must tend to zero as $n \to \infty$, preventing an optimal constant of $1/2$ in the limit. While this statement turns out to be false in the case $n = 1$, the following lemma will allow us to ignore that case by applying an $O(\delta)$ perturbation to $\mathbf{E}$ which makes $\mathcal{Q}(\pi_1(F_1), \pi_1(F_2)) = 0$.

Let Id denote the identity element of $O(d+1)$.

Lemma 4.11. *Let $d \ge 1$. Let $\mathbf{E}$ be as above, and let $\delta = dist(\mathbf{E}, \mathcal{O}(\mathbf{E}^*))$. Then there exists $Q \in O(d+1)$ satisfying $||Q - Id|| = O(\delta)$ such that the functions $\tilde{F}_j$ associated to the sets $\tilde{E}_j = Q(E_j)$ satisfy $\pi_1(\tilde{F}_1) = 0$.*

Proof. As $\tilde{E}_2$ and $\tilde{F}_2$ are absent from the conclusion, consider just E_1 and F_1, dropping the subscript to write them as E and F.

Let $Q \in O(d+1)$ and let $\tilde{F}$ be the function associated to the set $\tilde{E} = Q(E)$. Then,

$$\begin{aligned}\tilde{F}(\theta) - F(\theta) &= \int_{-1}^{1} (\mathbf{1}_{Q(E)} - \mathbf{1}_B)(\theta, t)(1-t^2)^{d/2} dt - \int_{-1}^{1} (\mathbf{1}_E - \mathbf{1}_B)(\theta, t)(1-t^2)^{d/2} dt \\ &= \int_{-1}^{1} (\mathbf{1}_{Q(E)} - \mathbf{1}_E)(\theta, t)(1-t^2)^{d/2} dt.\end{aligned}$$

Let $P \in \mathcal{H}_1$. Define $\langle F, P \rangle = \int_{S^{d-1}} F(\theta)P(\theta)d\theta$. Let $x = (\theta, t)$ be coordinates on S^d and $g(x) = g(\theta, t) = P(\theta)$. By linearity of the integral and an orthogonal change of coordinates $x \mapsto Qx$,

$$\begin{aligned}\langle F, P \rangle - \langle \tilde{F}, P \rangle &= \int_{S^d} (\mathbf{1}_{Q(E)} - \mathbf{1}_E)(x) g(x) d\sigma(x) \\ &= \int_{S^d} \mathbf{1}_E(x)[g \circ Q^{-1}(x) - g(x)] d\sigma(x).\end{aligned}$$

Consider F as an element of $\mathcal{H}_1^*$, the dual space of the vector space $\mathcal{H}_1$, by the linear mapping

$$\langle F, P \rangle = \int_{S^{d-1}} F(\theta)P(\theta)d\mu(\theta), \tag{4.34}$$

where $P \in \mathcal{H}_1$. Then

$$\begin{aligned}||\tilde{F} - F||_{\mathcal{H}_1^*} ||P||_\infty &\leq C||g \circ Q^{-1} - g||_{L^1(S^d)} \\ &\leq C||Q - \mathrm{Id}|| \cdot ||P||_\infty,\end{aligned}$$

where $|| \cdot ||$ is any norm on the finite-dimensional space of linear maps on $\mathbb{R}^{d+1}$. The second line is obtained from the first by splitting S^d into the small set where $|t| \approx 1$ and g is bounded in L^∞ norm, and the remaining set where g has bounded derivative. Thus, $||\tilde{F} - F||_{\mathcal{H}_1^*} = O(||Q - \mathrm{Id}||)$.

We may conclude the proof by a standard application of the Implicit Function Theorem, though it remains to be shown that $Q \mapsto \tilde{F}$ is locally surjective.

Identifying $O(d)$ with the subset of $O(d+1)$ of maps which preserve the coordinate t, we see that for any $Q \in O(d)$, $g \circ Q^{-1} - g \equiv 0$. If this were true for all $Q \in O(d+1)$, then no perturbation $\tilde{F}$ of F would satisfy $||\tilde{F}||_{\mathcal{H}_1^*} = 0$, as $\tilde{F}$ and F would always be equal. However, considering S^d as a subset of $\mathbb{R}^{n+1}$ with coordinates $(x_1, ..., x_{n+1})$, we pick d distinct choices of Q as rotations in the $x_i x_{n+1}$-plane for $1 \leq i \leq d$. These choices of Q determine d linearly independent values of $\tilde{F} - F$ in the d-dimensional space $\mathcal{H}_1^*$. In fact, one may see these elements of $\mathcal{H}_1^*$ have pairwise $O(\delta)$ inner product by testing them against the basis $\{x_1, ..., x_d\}$ for $\mathcal{H}_1$.

□

4.5 Completing the Proof

Given a pair of spherical harmonics $\mathbf{G} = (G_1, G_2)$ of equal degree n, define for all real s in a neighborhood of 0

$$E_j(s) := \{(\theta, t) : t \geq h_j - \varphi_j(\theta, s)\}, \tag{4.35}$$

where $\phi_j(\theta, s)$ is defined recursively via the equation $\int_{h_j-\varphi_j(\theta,s)}^{h_j}(1-t^2)^{d/2}dt = sG_j(\theta)$. Note that the functions $F_{j,s}$ associated to the $E_j(s)$ satisfies $F_{j,s} \equiv sG_j + O(s^2)$.

Lemma 4.12. *Let $d \geq 1$, $n \in \mathbb{N}$ and (e_1, e_2, a) be a strictly admissible triple. Then, uniformly for all pairs of spherical harmonics $\mathbf{G}$ of degree n satisfying $||\mathbf{G}|| = 1$, there exists $\eta > 0$ such that*

$$\mathcal{T}(\mathbf{E}(s)) = \mathcal{T}(\mathbf{E}^*) - \frac{1}{2}s^2\sum_{j=1}^{2}\gamma_j(1-h_j^2)^{-d/2}(||G_j||_{L^2}^2) + s^2\mathcal{Q}(\mathbf{G}) + O(s^3) \tag{4.36}$$

whenever $|s| \leq \eta$.

In making sense of the above lemma, note that the construction of $\mathbf{E}(s)$ depends solely on the values of h_j, which are in one-to-one correspondence with the values of $\sigma(E_j)$.

The proof of Lemma 4.12 is essentially the same as that of Proposition 4.8. The one difference is that the restrictions of the functions on S^d to a particular θ are in fact indicator functions of t. By considering only these particular choices of sets $\mathbf{E}$, one is able to reach a conclusion with equality in (4.36).

At this point, it is possible to conclude the proof of Theorem 4.2 in the case $d = 1$.

Proof of Theorem 4.2 when $d = 1$. Let $\tilde{E}_j$ be as in the conclusion of Lemma 4.11. Then, by the fact that $\pi_1(F_1) = 0$ and $\pi_0(F_1) = \pi_0(F_2) = 0$ (since $\int f_j = 0$), (4.33) holds trivially for $n = 0$ and $n = 1$. When $d = 1$, the sphere S^{d-1} consists of two points, so there are no spherical harmonics of degree greater than or equal to 2. Hence, (4.33) holds in all cases and the proof is complete. □

Now that the base case $d = 1$ is established, it is possible to apply Theorem 4.2 in dimension $d - 1$ as an inductive case to prove the theorem in dimension d. In particular, the inductive case will be used to prove the following lemma.

Lemma 4.13. *Let $d \geq 1$. There exists $c > 0$ such that for all pairs $\mathbf{G}$ of spherical harmonics such that $||\mathbf{G}|| = 1$ and $\deg(G_1) = \deg(G_2) = n \geq 2$,*

$$\mathcal{T}(\mathbf{E}(s)) \leq \mathcal{T}(\mathbf{E}^*) - cs^2 \tag{4.37}$$

for all $s \in \mathbb{R}$ sufficiently close to 0.

Combining Lemmas 4.12 and 4.13, one obtains (4.33) for all n, completing the proof of Theorem 4.2 for dimensions $d \geq 2$. Thus, it suffices to prove Lemma 4.13.

Proof. It suffices to prove the lemma for any single arbitrary such $\mathbf{G}$. Uniformity will follow by compactness of $\{\mathbf{G} : ||\mathbf{G}|| = 1, \deg(G_1) = \deg(G_2) = n\}$.

Of use is the following expression of $\mathcal{T}$ in terms of similar operators acting on sets of one lower dimension. Writing $x = (\theta_1, t_1), y = (\theta_2, t_2)$,

$$\begin{aligned}\iint_{(S^d)^2} &\mathbf{1}_{E_1}(x)\mathbf{1}_{E_2}(y)\mathbf{1}_I(x \cdot y)d\sigma(x)d\sigma(y)\\ &= \iint_{[-1,1]^2}\iint_{(S^{d-1})^2} \mathbf{1}_{E_1}(\theta_1, t_1)\mathbf{1}_{E_2}(\theta_2, t_2)\mathbf{1}_I\left(\sqrt{1-t_1^2}\sqrt{1-t_2^2}\theta_1\cdot\theta_2 + t_1t_2\right)\\ &\qquad\qquad \times d\mu(\theta_1)d\mu(\theta_2)(1-t_1^2)^{\frac{d}{2}}(1-t_2^2)^{\frac{d}{2}}dt_1dt_2 \end{aligned} \tag{4.38}$$

For fixed t_1, t_2, $\sqrt{1-t_1^2}\sqrt{1-t_2^2}\theta_1\cdot\theta_2 + t_1t_2 \geq a$ if and only if $\theta_1 \cdot \theta_2 \geq \frac{a - t_1t_2}{\sqrt{1-t_1^2}\sqrt{1-t_2^2}}$. Thus, we may write

$$\begin{aligned}&\mathcal{T}(\mathbf{E}(s), [a, 1]) =\\ &\iint_{[-1,1]^2} \mathcal{T}'\left(E_1(s; t_1), E_2(s; t_2), \left[\frac{a - t_1t_2}{\sqrt{1-t_1^2}\sqrt{1-t_2^2}}\right]\right)(1-t_1^2)^{d/2}(1-t_2^2)^{d/2}dt_1dt_2,\end{aligned} \tag{4.39}$$

where $E_j(s; t) = \{\theta : (\theta, t) \in E_j(s)\}$ and

$$\mathcal{T}'(A_1, A_2, I') := \iint_{S^{d-1}\times S^{d-1}} \mathbf{1}_{A_1}(\theta_1)\mathbf{1}_{A_2}(\theta_2)\mathbf{1}_{I'}(\theta_1\cdot\theta_2)d\mu(\theta_1)d\mu(\theta_2). \tag{4.40}$$

One may apply the spherical rearrangement inequality of Theorem 4.1 to the sets $E_j(s; t_j)$ on S^{d-1} for any fixed pair (t_1, t_2). It is not immediate that one may apply Theorem 4.2 for dimension $d - 1$, as there are additional hypotheses which must be satisfied. However, it will be shown that these hypotheses hold on a nontrivial subset S of $(h_1 - s, h_1 + s) \times (h_2 - s, h_2 + s)$ of size roughly proportional to s^2.

To construct S, observe that $V_j(s, t_j) := |E_j(s; t)|$ is a continuous function in t_j, taking value 0 at $t_j = h_j - s$ and $|S^{d-1}| = 1$ at $t_j = h_j + s$. Thus, we may choose

$$S = \{(t_1, t_2) \subset (h_1 - s, h_1 + s) \times (h_2 - s, h_2 + s) : \left(V_1(s, t_1), V_2(s, t_2), \frac{a - t_1 t_2}{\sqrt{1 - t_1^2}\sqrt{1 - t_2^2}} \right) \text{ is strictly admissible}\} \quad (4.41)$$

It immediately follows that S is an open subset of $(h_1 - s, h_1 + s) \times (h_2 - s, h_2 + s)$ which has nontrivial measure. The conclusion that the size of S is roughly proportional to s^2 follows from dilating it around (h_1, h_2), that $V_j(\lambda s, h_j + \lambda a_j) = V_j(s, h_j + a_j) + O(s^2)$, and the fact that $\frac{a - t_1 t_2}{\sqrt{1 - t_1^2}\sqrt{1 - t_2^2}}$ is locally constant near $t_j = h_j$.

Next, to get the desired gain from the inductive hypothesis, one must show that $\mathrm{dist}(E_j(s; t_j), \mathcal{O}(E_j(s; t_j))) > 0$. This follows from the fact that for a degree 2 or greater spherical harmonic G, the sets $\{\theta : G(\theta) > \alpha\}$ are not spherical caps for α close to 0.

□

Bibliography

[1] K. I. Babenko. An inequality in the theory of Fourier integrals. *Izv. Akad. Nauk SSSR Ser. Mat.*, 25:531–542, 1961.

[2] Albert Baernstein, II. Convolution and rearrangement on the circle. *Complex Variables Theory Appl.*, 12(1-4):33–37, 1989.

[3] Albert Baernstein, II and B. A. Taylor. Spherical rearrangements, subharmonic functions, and *-functions in n-space. *Duke Math. J.*, 43(2):245–268, 1976.

[4] William Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[5] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao. The Brascamp-Lieb inequalities: finiteness, structure and extremals. *Geom. Funct. Anal.*, 17(5):1343–1415, 2008.

[6] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao. Finite bounds for Hölder-Brascamp-Lieb multilinear inequalities. *Math. Res. Lett.*, 17(4):647–666, 2010.

[7] Gabriele Bianchi and Henrik Egnell. A note on the Sobolev inequality. *J. Funct. Anal.*, 100(1):18–24, 1991.

[8] Garrett Birkhoff. *Lattice theory.* Third edition. American Mathematical Society Colloquium Publications, Vol. XXV. American Mathematical Society, Providence, R.I., 1967.

[9] Herm Jan Brascamp and Elliott H. Lieb. Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Math.*, 20(2):151–173, 1976.

[10] Almut Burchard. Cases of equality in the Riesz rearrangement inequality. *Ann. of Math. (2)*, 143(3):499–527, 1996.

[11] E. A. Carlen, E. H. Lieb, and M. Loss. A sharp analog of Young's inequality on S^N and related entropy inequalities. *J. Geom. Anal.*, 14(3):487–520, 2004.

[12] Michael Christ. Near-extremizers of Young's Inequality for $\mathbb{R}^d$. *ArXiv e-prints*, December 2011.

[13] Michael Christ. Near equality in the Riesz-Sobolev inequality. *arXiv e-prints*, page arXiv:1309.5856, Sep 2013.

[14] Michael Christ. A sharpened Hausdorff-Young inequality. *arXiv e-prints*, page arXiv:1406.1210, Jun 2014.

[15] Michael Christ. A sharpened Riesz-Sobolev inequality. *arXiv e-prints*, page arXiv:1706.02007, Jun 2017.

[16] Michael Christ. On Young's inequality for Heisenberg groups. *arXiv e-prints*, page arXiv:1706.02005, Jun 2017.

[17] Michael Christ. Young's Inequality Sharpened. *preprint, submitted*, 2018.

[18] Michael Christ and Marina Iliopoulou. Inequalities of Riesz-Sobolev type for compact connected Abelian groups. *arXiv e-prints*, page arXiv:1808.08368, Aug 2018.

[19] Michael Christ and Kevin O'Neill. Maximizers of Rogers-Brascamp-Lieb-Luttinger functionals in higher dimensions. *arXiv e-prints*, page arXiv:1712.00109, Nov 2017.

[20] J. A. Crowe, J. A. Zweibel, and P. C. Rosenbloom. Rearrangements of functions. *J. Funct. Anal.*, 66(3):432–438, 1986.

[21] James Demmel and Alex Rusciano. Parallelepipeds obtaining HBL lower bounds. *arXiv e-prints*, page arXiv:1611.05944, Nov 2016.

[22] Thomas S. Ferguson. Linear programming: A concise introduction. https://www.math.ucla.edu/ tom/LP.pdf.

[23] P. Ivanisvili and A. Volberg. Hessian of Bellman functions and uniqueness of the Brascamp-Lieb inequality. *J. Lond. Math. Soc. (2)*, 92(3):657–674, 2015.

[24] Abel Klein and Bernard Russo. Sharp inequalities for Weyl operators and Heisenberg groups. *Math. Ann.*, 235(2):175–194, 1978.

[25] Elliott H. Lieb. Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities. *Ann. of Math. (2)*, 118(2):349–374, 1983.

[26] Elliott H. Lieb. Gaussian Kernels have only Gaussian Maximizers. *Invent. Math.*, 102:179–208, 1990.

[27] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[28] Frédéric Riesz. Sur Une Inegalite Integarale. *J. London Math. Soc.*, 5(3):162–168, 1930.

[29] Stanislaw Saks. *Theory of the Integral*. Second Revised Edition. Hafner Publishing Company, New York, 1937.

[30] S.L. Sobolev. On a theorem of functional analysis. *Mat. Sb. (N.S.)*, (4), 1938.

[31] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993.