ANALYSIS & PDE

Volume 5 No. 2 2012

MICHAEL CHRIST AND SHUANGLIN SHAO

EXISTENCE OF EXTREMALS FOR A FOURIER RESTRICTION INEQUALITY

mathematical sciences publishers

ANALYSIS AND PDE
Vol. 5, No. 2, 2012
dx.doi.org/10.2140/apde.2012.5.261

EXISTENCE OF EXTREMALS FOR A FOURIER RESTRICTION INEQUALITY

MICHAEL CHRIST AND SHUANGLIN SHAO

The adjoint Fourier restriction inequality of Tomas and Stein states that the mapping $f \mapsto \widehat{f\sigma}$ is bounded from $L^2(\mathbb{S}^2)$ to $L^4(\mathbb{R}^3)$. We prove that there exist functions that extremize this inequality, and that any extremizing sequence of nonnegative functions has a subsequence that converges to an extremizer.

1. Introduction

Let $\mathbb{S}^2$ denote the unit sphere in $\mathbb{R}^3$, equipped with surface measure σ. The adjoint Fourier restriction inequality of Tomas and Stein, for $\mathbb{S}^2$, states that there exists $C < \infty$ such that

$$\|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \leq C\|f\|_{L^2(\mathbb{S}^2,\sigma)} \tag{1-1}$$

Christ was supported in part by NSF grant DMS-0901569. Shao was supported by the National Science Foundation under agreement DMS-0635607.
MSC2000: 42A38.
Keywords: extremals, adjoint Fourier restriction inequality.

for all $f \in L^2(\mathbb{S}^2)$. With the Fourier transform defined to be $\hat{g}(\xi) = \int e^{-ix\cdot\xi} g(x)\,dx$, denote by

$$\mathcal{R} = \sup_{0 \neq f \in L^2(\mathbb{S}^2)} \|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \Big/ \|f\|_{L^2(\mathbb{S}^2,\sigma)}$$

the optimal constant in the inequality (1-1).

Definition 1.1. An extremizing sequence for the inequality (1-1) is a sequence $\{f_\nu\}$ of functions in $L^2(\mathbb{S}^2)$ satisfying $\|f_\nu\|_2 \le 1$, such that $\|\widehat{f_\nu\sigma}\|_{L^4(\mathbb{R}^3)} \to \mathcal{R}$ as $\nu \to \infty$.

An extremizer for the inequality (1-1) is a function $f \neq 0$ that satisfies $\|\widehat{f\sigma}\|_4 = \mathcal{R}\|f\|_2$.

The main result of this paper is this:

Theorem 1.2. *There exists an extremizer in $L^2(\mathbb{S}^2)$ for the inequality* (1-1).

The inequality dual to (1-1) is $\|\hat{h}\|_{L^2(\mathbb{S}^2,\sigma)} \le C\|h\|_{L^{4/3}(\mathbb{R}^3)}$. If f extremizes (1-1), then $\widehat{f\sigma} \cdot |\widehat{f\sigma}|^2$ extremizes the dual inequality.

Our inequality is one of endpoint type. That is, it becomes false if either of the exponents $2, 4$ is decreased. An analogue of Theorem 1.2 has more recently been obtained by Fanelli, Vega, and Visciglia [Fanelli et al. 2011], for adjoint restriction inequalities not of endpoint type.

Definition 1.3. A sequence of functions in $L^2(\mathbb{S}^2)$ is precompact if any subsequence has a sub-sub-sequence that is Cauchy in $L^2(\mathbb{S}^2)$.

Nonnegative functions play a special role in our analysis, because

$$\|\widehat{|f|\sigma}\|_4 \ge \|\widehat{f\sigma}\|_4 \quad \text{for all } f \in L^2(\mathbb{S}^2).$$

Therefore if $\{f_\nu\}$ is an extremizing sequence, so is $\{|f_\nu|\}$. Any limit, in the L^2 norm, of an extremizing sequence is of course an extremizer. Thus the following implies Theorem 1.2.

Theorem 1.4. *Any extremizing sequence of nonnegative functions in $L^2(\mathbb{S}^2)$ for the inequality* (1-1) *is precompact.*

In particular, the set of all nonnegative extremizers is itself compact. We do not know whether nonnegative extremizers are unique modulo rotations of $\mathbb{S}^2$ and multiplication by constants. They do possess the following symmetry, which will be useful in our analysis.

Theorem 1.5. *Every extremizer satisfies $|f(-x)| = |f(x)|$ for almost every $x \in \mathbb{S}^2$.*

Proposition 2.7 below states that more generally, the quantity $\|\widehat{f\sigma}\|_4$ never decreases under L^2 norm-preserving symmetrization of f with respect to the map $x \mapsto -x$.

For complex-valued extremizers and near extremizers, the situation regarding precompactness of extremizing sequences is different, due to the presence of a noncompact group of symmetries of the inequality. For $\xi \in \mathbb{C}^3$, define $e_\xi(x) = e^{x\cdot\xi}$. Then $\|\widehat{fe_{i\xi}\sigma}\|_4 = \|\widehat{f\sigma}\|_4$ for arbitrary $\xi \in \mathbb{R}^3$, where $f \in L^2(\mathbb{S}^2)$. Consequently complex-valued extremizing sequences need not be precompact. However, we show in a sequel [Christ and Shao 2012] that this simple obstruction is the only one; if $\{f_\nu\}$ is any complex-valued extremizing sequence, then there exists a sequence $\{\xi_\nu\} \subset \mathbb{R}^3$ such that $e^{-ix\cdot\xi_\nu} f_\nu(x)$ is precompact.

The symmetries $f \mapsto f \cdot e^{ix\cdot\xi}$ merit further discussion. Matters are clearer for the paraboloid $\mathbb{P}^2 = \{(y_1, y_2, y_3) : y_3 = \frac{1}{2}y_1^2 + \frac{1}{2}y_2^2\}$ than for $\mathbb{S}^2$. For $\mathbb{P}^2$, the analogues of these unimodular exponentials are *quadratic* exponentials $e^{ix\cdot\eta + i\tau|x|^2}$ with $(\eta, \tau) \in \mathbb{R}^{2+1}$; compare with $\mathbb{S}^2$, where $\xi \in \mathbb{R}^3$ also ranges over a three-dimensional space. To see the analogy, consider a small neighborhood of $(0, 0, 1) \in \mathbb{S}^2$, equipped with coordinates $x' \in \mathbb{R}^2$ such that $x = (x', (1 - |x'|^2)^{1/2})$. Then for $\xi = (0, 0, \lambda)$, we have $e^{ix\cdot\xi} = \exp(i\lambda(1 - \frac{1}{2}|x'|^2 + O(|x'|^4))$ for small x'; thus for small x' one has essentially quadratic oscillation. The presence of these symmetries among the extremizers for $\mathbb{P}^2$ implies that, in the language of concentration compactness theory [Kunze 2003], an extremizer f can be tight at a scale r, and $\hat{f}$ can simultaneously be tight at a scale $\hat{r}$, with the product $r \cdot \hat{r}$ arbitrarily large.

Define

$$\mathbf{S} := \sup_{0 \neq f \in L^2(\mathbb{S}^2, \sigma)} \frac{\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)}^{1/2}}{\|f\|_{L^2(\mathbb{S}^2,\sigma)}}.$$

Then $\mathcal{R} = (2\pi)^{3/4}\mathbf{S}$ by Plancherel's theorem and the connection between the Fourier transform and convolution.

$\mathbf{S}$ is the supremum of a functional, whose critical points are characterized by the generalized Euler–Lagrange equation

$$(f\sigma * f\sigma * \tilde{f}\sigma)|_{\mathbb{S}^2} = \lambda f \quad \text{almost everywhere on } \mathbb{S}^2, \tag{1-2}$$

where $\tilde{f}(x) = \overline{f(-x)}$ and λ is a Lagrange multiplier determined by f. This follows from a routine variational argument; see for instance [Christ and Quilodrán 2010], where more general results of this type are justified. Equation (1-2) will be used in a forthcoming paper [Christ and Shao 2012] to prove that all critical points are infinitely differentiable. By taking the $L^2(\mathbb{S}^2)$ inner product of both sides with f, one obtains an alternative characterization of extremizers.

Proposition 1.6. *A complex-valued function $f \in L^2(\mathbb{S}^2)$ is an extremizer if and only if*

$$(f\sigma * f\sigma * \tilde{f}\sigma)|_{\mathbb{S}^2} = \mathbf{S}^4 \|f\|_2^2 f \quad \textit{almost everywhere on } \mathbb{S}^2,$$

where $\tilde{f}(x) = \overline{f(-x)}$.

Since the numerical value of $\mathbf{S}$ has not been determined, this equation is not entirely explicit and provides only a negative test for extremizers.

Fundamental questions remain open, among them these:

Questions 1.7. Are extremizers unique modulo rotations and multiplication by constants? Are constant functions extremizers?

In this context, it is interesting to observe that constant functions are *local* maxima. Let $\mathbf{1}$ denote the constant function $f(x) \equiv 1$.

Theorem 1.8. *There exists $\delta > 0$ such that whenever $\|f - \mathbf{1}\|_{L^2(\mathbb{S}^2)} < \delta$,*

$$\frac{\|\widehat{f\sigma}\|_4^4}{\|f\|_2^4} \leq \frac{\|\widehat{\sigma}\|_4^4}{\|\mathbf{1}\|_2^4},$$

with equality only if f is constant.

Let $\mathbb{P}^2$ be the paraboloid introduced above. Let σ_P be the measure $d\sigma_P = dx_1\,dx_2$ on $\mathbb{P}^2$.[1] Then the mapping $f \mapsto \widehat{f\sigma_P}$ is likewise bounded from $L^2(\mathbb{P}^2, \sigma_P)$ to $L^4(\mathbb{R}^3)$. Denote by $\mathcal{R}_{\mathbb{P}^2}$ the optimal constant in the inequality

$$\|\widehat{f\sigma_P}\|_{L^4(\mathbb{R}^3)} \le \mathcal{R}_{\mathbb{P}^2}\|f\|_{L^2(\mathbb{P}^2,\sigma_P)}. \tag{1-3}$$

Foschi [2007] has proved that extremals exist for this inequality, and moreover, that every radial Gaussian $f(x', x_3) = e^{-c|x'|^2}$ is an extremal, where $x' = (x_1, x_2)$, and that $\mathcal{R}_{\mathbb{P}^2} = 2^{3/4}\pi$. Alternative proofs were given by Hundertmark and Zharnitsky [2006] and by Bennett, Bez, Carbery, and Hundertmark [Bennett et al. 2009]. The simple relation $\mathcal{R} \ge \mathcal{R}_{\mathbb{P}^2}$ is of significance for our discussion. This relation follows from examination of a suitable sequence of trial functions f_ν, such that $f_\nu(x)^2$ converges weakly to a Dirac mass on $\mathbb{S}^2$, and f_ν is approximately a Gaussian in suitably rescaled coordinates, depending on ν. It is essential for this comparison that $\mathbb{P}^2$ has the same curvature at 0 as $\mathbb{S}^2$, which explains the factors of $\frac{1}{2}$ in the definition of $\mathbb{P}^2$.

The first author to discuss existence of extremizers for Strichartz/Fourier restriction inequalities was apparently Kunze [2003], who proved the existence of extremizers for the parabola in $\mathbb{R}^2$, and showed that (in our notation) any nonnegative extremizing sequence is precompact modulo the action of the natural symmetry group of the inequality. Several papers have subsequently dealt with related problems, in some cases determining all extremizers explicitly [Foschi 2007; Hundertmark and Zharnitsky 2006; Bennett et al. 2009; Carneiro 2009], in other cases merely proving existence [Shao 2009]. A powerful result [Shao 2009] that leads easily to existence of extremizers is the profile decomposition; see [Bégout and Vargas 2007]. Of these works, the one most closely related to ours is that of Kunze. One difficulty that we face is the lack of exact scaling symmetries. In some facets of the analysis this is merely a technical obstacle, but it is bound up with the most essential obstacle, which is the possibility that the optimal constant might be achieved only in a limit where $|f_\nu|^2$ tends to a Dirac mass, or a sum of two Dirac masses.

Our analysis follows the general concentration compactness framework developed by Lions [1984a; 1984b; 1985a; 1985b]. We have elected to make the exposition self-contained in this respect, not drawing on that theory; to do so would apparently not dramatically shorten the exposition, since most of our labor is lavished on specific issues raised by the character of a particular nonlocal operator.

Existence of extremals for another scale-invariant convolution inequality in which curvature plays an essential role, as it does here, was proved in [Christ 2011a]. There the underlying geometry is more subtle, but the operator analyzed is merely linear, while the analysis of this paper is bilinear. Despite differences in details, that analysis and the method of this paper have much in common. The role of an inequality of Moyua, Vargas, and Vega [Moyua et al. 1999] used here was played in [Christ 2011a] by [Christ 2011b].

[1] See [Christ 2011a] for a brief discussion of the naturality of this measure from a geometric perspective.

2. Outline of the proof and definitions

The following overview of the proof includes notations, definitions, and statements of intermediate results that are not repeated subsequently, and thus is an integral part of the presentation.

Step 1. The first step is quite simple, but in it a critical distinction appears between our problem for $\mathbb{S}^2$, and for higher-dimensional spheres. The inequality $\|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \le \mathcal{R}\|f\|_{L^2(\mathbb{S}^2,\sigma)}$ is equivalent, by squaring and Plancherel's theorem, to

$$\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)} \le \mathbf{S}^2\|f\|^2_{L^2(\mathbb{S}^2)}, \tag{2-1}$$

where

$$\mathcal{R} = (2\pi)^{3/4}\mathbf{S}$$

and $*$ denotes convolution of measures. This has been exploited in [Kunze 2003; Foschi 2007; Hundertmark and Zharnitsky 2006; Bennett et al. 2009]. In higher dimensions, the exponent 4 is replaced by an exponent that is no longer an even integer, and no such equivalence is available.

Now the pointwise inequality $|f\sigma * f\sigma| \le |f|\sigma * |f|\sigma$, the relation $\widehat{\mu * \nu} = \hat{\mu}\hat{\nu}$, and Plancherel's theorem imply this:

Lemma 2.1. *For any complex-valued function $f \in L^2(\mathbb{S}^2)$,*

$$\|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \le \|\widehat{|f|\sigma}\|_{L^4(\mathbb{R}^3)}.$$

Therefore if f is an extremizer for inequality (1-1), *then so is $|f|$; if $\{f_\nu\}$ is an extremizing sequence, so is $\{|f_\nu|\}$.*

This permits us to work with nonnegative functions throughout the analysis. For much of our analysis this makes no difference, but nonnegativity will be useful in Step 7, allowing an elementary approach to a step whose analogue in higher dimensions seems to require more sophisticated techniques.

Step 2. A potential obstruction to the existence of extremizers, and certainly to the precompactness of arbitrary extremizing sequences, is the possibility that for an extremizing sequence satisfying $\|f_\nu\|_2 = 1$, $|f_\nu|^2$ could conceivably converge weakly to a Dirac mass at a point of $\mathbb{S}^2$. Straightforward analysis of a sequence $\{f_\nu\}$ chosen so that $|f_\nu|^2$ converges in this way, disregarding the question of whether $\{f_\nu\}$ is extremizing, reveals that $\mathcal{R} \ge \mathcal{R}_{\mathbb{P}^2}$; see Lemma 3.1. Now if $\mathcal{R}$ were to equal $\mathcal{R}_{\mathbb{P}^2}$, any such sequence would be extremizing, yet would not be precompact. Therefore an unavoidable step in our analysis is to demonstrate a strict inequality $\mathcal{R} > \mathcal{R}_{\mathbb{P}^2}$.

In fact, as will be explained below, this is true in two distinct ways. The more superficial is this:

Lemma 2.2. *Let $g \in L^2(\mathbb{S}^2)$ be supported in $\{x \in \mathbb{S}^2 : x_3 > \frac{1}{2}\}$. Define $f(x) = 2^{-1/2}g(x) + 2^{-1/2}\overline{g(-x)}$. Then $\|f\|_2 = \|g\|_2$, and*

$$\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)} = (3/2)^{1/2}\|g\sigma * g\sigma\|_{L^2(\mathbb{R}^3)}.$$

Define the optimal constant in the corresponding inequality for the paraboloid to be

$$\mathbf{P} = \sup_{0 \neq g \in L^2(\mathbb{P}^2, \sigma_P)} \frac{\|g\sigma_P * g\sigma_P\|_{L^2(\mathbb{R}^3)}^{1/2}}{\|g\|_{L^2(\mathbb{P}^2, \sigma_P)}}.$$

By Lemma 2.2, the optimal constants for $\mathbb{S}^2$ and $\mathbb{P}^2$ satisfy the following.

Corollary 2.3. $$\mathbf{S} \geq (3/2)^{1/4}\mathbf{P}.$$

Step 3. The simplest possibility left open by Step 2 is that an extremizing sequence might concentrate at a pair of antipodal points, that is, $|f_\nu|^2$ might converge weakly to a linear combination of two Dirac masses, at antipodal points z and $-z$. This scenario is indeed the crux of the problem. The crucial ingredient in excluding it is an improved inequality $\mathbf{S} > (3/2)^{1/4}\mathbf{P}$. We will give two independent proofs of this inequality. The first gives a precise improvement:

Lemma 2.4. $$\mathbf{S} \geq 2^{1/4}\mathbf{P}.$$

Equivalently, $\mathcal{R} \geq 2^{1/4}\mathcal{R}_{\mathbb{P}^2}$. This is proved by an exact computation of $\|f\sigma * f\sigma\|_2$ for $f \equiv 1$. We do not know whether constant functions are in fact extremal for (1-1), or equivalently, whether $\mathbf{S} = 2^{1/4}\mathbf{P}$. Constants are indeed critical points of the associated functional, and thus satisfy a (possibly) modified Euler–Lagrange equation (1-2), in which $\mathbf{S}$ is replaced by $2^{1/4}\mathbf{P}$.

An alternative proof that $\mathbf{S} > (3/2)^{1/4}\mathbf{P}$, along perturbative lines, is given in Section 17.

Step 4. Definition 2.5. A complex-valued function $f \in L^2(\mathbb{S}^2)$ is said to be even if $f(-x) = \overline{f(x)}$ for almost every $x \in \mathbb{S}^2$.

We will be working almost exclusively with nonnegative functions, for which this condition becomes $f(-x) \equiv f(x)$.

Definition 2.6. Let $f \in L^2(\mathbb{S}^2)$ be nonnegative. The antipodally symmetric rearrangement $f_\star$ is the unique nonnegative element of $L^2(\mathbb{S}^2)$ that satisfies

$$\begin{aligned} f_\star(-x) &= f_\star(x) && \text{for all } x \in \mathbb{S}^2, \\ f_\star(x)^2 + f_\star(-x)^2 &= f(x)^2 + f(-x)^2 && \text{for all } x \in \mathbb{S}^2. \end{aligned}$$

In other words, $f_\star(x) = \sqrt{(f(x)^2 + f(-x)^2)/2}$ for all $x \in \mathbb{S}^2$.

Proposition 2.7. *For any nonnegative $f \in L^2(\mathbb{S}^2)$,*

$$\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)} \leq \|f_\star\sigma * f_\star\sigma\|_{L^2(\mathbb{R}^3)},$$

with strict inequality unless $f = f_\star$ almost everywhere. Consequently any extremizer for the inequality (1-1) *satisfies $|f(-x)| = |f(x)|$ for almost every $x \in \mathbb{S}^2$.*

An equivalent formulation is that $\|\widehat{f\sigma}\|_4 \leq \|\widehat{f_\star\sigma}\|_4$.

This allows us to restrict attention from nonnegative functions to even nonnegative functions throughout the discussion. This simplification is more convenient than essential.

Step 5. A first key step towards gaining control of near-extremals has already been essentially accomplished by Moyua, Vargas, and Vega [Moyua et al. 1999].

Definition 2.8. The cap $\mathscr{C} = \mathscr{C}(z, r)$ with center $z \in \mathbb{S}^2$ and radius $r \in (0, 1]$ is the set of all points $y \in \mathbb{S}^2$ that lie in the same hemisphere, centered at z, as z itself, and that satisfy $|\pi_{H_z}(y)| < r$, where the subspace $H_z \subset \mathbb{R}^3$ is the orthogonal complement of z and π_{H_z} denotes the orthogonal projection onto H_z.

Lemma 2.9. *For any $\delta > 0$ there exist $C_\delta < \infty$ and $\eta_\delta > 0$ with the following property. If $f \in L^2(\mathbb{S}^2)$ satisfies $\|f\sigma * f\sigma\|_2 \ge \delta^2 \mathbf{S}^2 \|f\|_2^2$, then there exist a decomposition $f = g + h$ and a cap $\mathscr{C}$ satisfying*

$$
\begin{aligned}
&0 \le |g|, |h| \le |f|,\\
&g \text{ and } h \text{ have disjoint supports},\\
&|g(x)| \le C_\delta \|f\|_2 |\mathscr{C}|^{-1/2} \chi_{\mathscr{C}}(x) \quad \text{for all } x,\\
&\|g\|_2 \ge \eta_\delta \|f\|_2.
\end{aligned}
$$

The first conclusion is of course redundant. If $f \ge 0$ then it follows that $g, h \ge 0$ almost everywhere.

Lemma 2.9 is a corollary [Moyua et al. 1999, Theorem 4.2]. It can also be proved via arguments closely related to those in [Christ 2011b].

Step 6. This step is related to the techniques used in [Christ 2011a].

Definition 2.10. Let $\mathscr{C} = \mathscr{C}(z, r)$ be a cap. For $z \in \mathbb{S}^2$, define $\psi_z(x) = r^{-1} L(\pi_{H_z}(x))$ for x in the hemisphere $\{x : x \cdot z > 0\}$, where π_{H_z} is the orthogonal projection onto H_z, and $L = L_z : H_z \to \mathbb{R}^2$ is an arbitrary linear isometry. The rescaling map associated with $\mathscr{C}$ is defined by $\phi_{\mathscr{C}(z,r)} = \psi_z^{-1}$.

The map $\phi_{\mathscr{C}(z,r)}$ is a bijection from $B(0, r^{-1}) \subset \mathbb{R}^2$ to the indicated hemisphere. For $z = (0, 0, 1)$, $\phi_{\mathscr{C}(z,r)}(y_1, y_2) = (ry_1, ry_2, (1 - r^2|y|^2)^{1/2})$ for $y \in B(0, r^{-1})$.

Definition 2.11. Let $\mathscr{C} = \mathscr{C}(z, r)$ be a cap. For $f \in L^2(\mathbb{S}^2)$, define the pullback of f by

$$\phi_{\mathscr{C}}^* f(y) = r \cdot (f \circ \phi_{\mathscr{C}})(y).$$

These pullbacks preserve norms up to uniformly bounded factors provided that $r \le r_0 < 1$; we have $\|\phi_{\mathscr{C}}^* f\|_{L^2(\mathbb{R}^2)} \asymp \|f\|_{L^2(\mathbb{S}^2,\sigma)}$, with the ratio of these norms bounded above and below by positive, finite constants, uniformly in f, r, z. For the sake of definiteness only, we will sometimes set $r_0 = \frac{1}{2}$.

Definition 2.12. Let $\Theta : [1, \infty) \to (0, \infty)$ satisfy $\Theta(R) \to 0$ as $R \to \infty$, and $\mathscr{C} = \mathscr{C}(z, r) \subset \mathbb{S}^2$ be a cap of radius r and center z. A function $f \in L^2(\mathbb{S}^2)$ is said to be upper normalized, with gauge function Θ, with respect to $\mathscr{C}$, if

$$\|f\|_2 \le C < \infty, \tag{2-2}$$

$$\int_{|f(x)| \ge Rr^{-1}} |f(x)|^2 \, d\sigma(x) \le \Theta(R) \qquad \text{for all } R \ge 1, \tag{2-3}$$

$$\int_{|x-z| \ge Rr} |f(x)|^2 \, d\sigma(x) \le \Theta(R) \qquad \text{for all } R \ge 1. \tag{2-4}$$

An even function f is said to be upper even-normalized with respect to $\Theta, \mathscr{C}(z,r)$ if, when f is decomposed as $f = f_+ + f_-$, where f_+ is the restriction of f to the hemisphere $\{x \in \mathbb{S}^2 : x \cdot z > 0\}$, the summand f_+ is upper normalized with respect to $\Theta, \mathscr{C}(z,r)$.

A function $f \in L^2(\mathbb{R}^2)$ is said to be upper normalized with respect to the unit ball in $\mathbb{R}^2$ if $\|f\|_2 \le C < \infty$, $\int_{|f(x)|\ge R} |f(x)|^2\,dx \le \Theta(R)$ for all $R \ge 1$, and $\int_{|x|\ge R} |f(x)|^2\,dx \le \Theta(R)$ for all $R \ge 1$.

For an even function f, we have $f_-(x) \equiv \overline{f_+(-x)}$, for almost every $x \in \mathbb{S}^2$. We will usually omit the phrase "with gauge function Θ", and will say that a function is upper normalized if it satisfies the required inequalities with respect to some appropriate function Θ which has been, in principle, specified earlier in the discussion.

Definition 2.13. A nonzero function $f \in L^2(\mathbb{S}^2)$ is said to be δ-nearly extremal for the inequality (2-1) if

$$\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)} \ge (1-\delta)^2 \mathbf{S}^2 \|f\|_2^2.$$

Proposition 2.14. *There exists a function $\Theta : [1,\infty) \to (0,\infty)$ satisfying $\Theta(R) \to 0$ as $R \to \infty$ with the following property. For any $\varepsilon > 0$ there exists $\delta > 0$ such that any nonnegative even function $f \in L^2(\mathbb{S}^2)$ satisfying $\|f\|_2 = 1$ that is δ-nearly extremal may be decomposed as $f = F + G$, where F and G are even and nonnegative with disjoint supports, $\|G\|_2 < \varepsilon$, and there exists a cap $\mathscr{C}$ such that F is upper even-normalized with respect to $\mathscr{C}$.*

The proof is a largely formal argument that rests on two inputs: Lemma 2.9, and the observation that $\|\chi_{\mathscr{C}}\sigma * \chi_{\mathscr{C}'}\sigma\|_2 \ll |\mathscr{C}|^{1/2}|\mathscr{C}'|^{1/2}$ for two caps $\mathscr{C}$ and $\mathscr{C}'$, unless they have comparable radii and nearby centers.

Step 7. In this step we establish *a priori* bounds for extremizing sequences, which include a limited but uniform smoothness after suitable rescaling. Step 7 and the closely related Step 9 are the only ones that require nonnegative extremizing sequences.

Proposition 2.15. *Let $\{f_\nu\} \subset L^2(\mathbb{S}^2)$ be an extremizing sequence of nonnegative even functions for the inequality* (2-1), *satisfying $\|f_\nu\|_2 \equiv 1$. Suppose that each f_ν is upper even-normalized with respect to a cap $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$, with constants uniform in ν. Assume that $\lim_{\nu\to\infty} r_\nu = 0$. Then for any $\varepsilon > 0$ there exists $C_\varepsilon < \infty$ such that each $\phi_\nu^*(f_\nu)$ may be decomposed as $\phi_\nu^*(f_\nu) = G_\nu + H_\nu$ where*

$$\|H_\nu\|_2 < \varepsilon,$$

$$G_\nu \textit{ is supported where } |x| \le C_\varepsilon,$$

$$\|G_\nu\|_{C^1} \le C_\varepsilon.$$

Here $\phi_\nu^* = \phi_{\mathscr{C}_\nu}^*$.

Proposition 2.15 expresses a weak form of equicontinuity, after rescaling. In outline: If $g \in L^2(\mathbb{R}^2)$ satisfies $\|g\|_2 \sim 1$, if g is upper normalized with respect to the unit ball, and if g is nonnegative, then $\int_{|\xi|\le 1} |\hat{g}(\xi)|^2\,d\xi$ is bounded below by a universal strictly positive constant. If the conclusions of the proposition were to fail, then $g_\nu = \phi_\nu^*(f_\nu)$ would have to satisfy $\int_{|\xi|} \ge \Lambda_\nu |\widehat{g_\nu}(\xi)|^2\,d\xi \ge \eta > 0$, with

$\limsup \Lambda_\nu = \infty$. Thus in an appropriately rescaled sense, for some subsequence, f_ν would be a superposition of a slowly varying part and a highly oscillatory part, with perhaps some intermediate portion of arbitrarily small norm for large ν. For the bilinear expression $f\sigma * f\sigma$, we show that the cross term resulting from the high and low frequency parts is small, and that this contradicts extremality.

Step 8. Proposition 2.16. *Let* $\{f_\nu\} \subset L^2(\mathbb{S}^2)$ *be an extremizing sequence of nonnegative even functions for the inequality* (2-1), *satisfying* $\|f_\nu\|_2 \equiv 1$. *Suppose that each* f_ν *is upper even-normalized with respect to a cap* $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$, *with constants uniform in* ν. *Then* $\inf_\nu r_\nu > 0$.

Thus the situation considered in the hypotheses of Proposition 2.15 cannot arise. The proof of Proposition 2.16 proceeds by contradiction. One can assume that $r_\nu \to 0$. A natural rescaling and transference procedure constructs a corresponding sequence of functions $\{f_\nu^+\}$ on $\mathbb{P}^2$, which possesses a weak form of equicontinuity, as a consequence of Proposition 2.15. In coordinates rescaled according to r_ν, each f_ν^+ is acted upon by an adjoint Fourier restriction operator associated to a hypersurface that depends on r_ν, and that approaches $\mathbb{P}^2$ as $r_\nu \to 0$. The weak equicontinuity of $\{f_\nu^+\}$, combined with the convergence of these hypersurfaces, can be used to construct a new sequence $F_\nu \in L^2(\mathbb{P}^2)$ that satisfies $\limsup_{\nu\to\infty} \|\widehat{F_\nu \sigma_P}\|_4 / \|F_\nu\|_2 \geq (3/2)^{-1/4} \lim_{\nu\to\infty} \|\widehat{f_\nu\sigma}\|_4/\|f_\nu\|_2$. It follows that $\mathscr{R}_{\mathbb{P}^2} \geq (3/2)^{-1/4}\mathscr{R}$. But this contradicts the inequality $\mathscr{R} \geq 2^{1/4}\mathscr{R}_{\mathbb{P}^2}$ of Step 3.

Step 9. The following variant of Proposition 2.15 is proved by essentially the same reasoning, with one small modification.

Proposition 2.17. *Let* $\{f_\nu\} \subset L^2(\mathbb{S}^2)$ *be an extremizing sequence of nonnegative even functions for the inequality* (2-1), *satisfying* $\|f_\nu\|_2 \equiv 1$. *Suppose that each* f_ν *is upper even-normalized with respect to a cap* $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$, *with constants uniform in* ν. *Let* $\rho > 0$, *and suppose that* $r_\nu \geq \rho$ *for every* ν. *Then after passing to some subsequence of* $\{r_\nu\}$, *each* f_ν *may be decomposed as* $f_\nu = g_\nu + h_\nu$, *where* $\|h_\nu\|_2 < \varepsilon$ *and* $\|g_\nu\|_{C^1} \leq C_{\varepsilon,\rho}$, *where* $C_{\varepsilon,\rho}$ *depends only on* ε, ρ, *not on* ν.

An application of Rellich's lemma yields precompactness:

Corollary 2.18. *Let* $\{f_\nu\} \subset L^2(\mathbb{S}^2)$ *be an extremizing sequence of even nonnegative functions for the inequality* (2-1), *which are upper even-normalized with respect to a sequence of caps* $\{\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)\}$. *Then* $\{f_\nu\}$ *is precompact in* $L^2(\mathbb{S}^2)$.

Conclusion. Extremizing sequences exist. We have shown that there exists an extremizing sequence that consists of even, nonnegative functions. Such a sequence is upper even-normalized with respect to a sequence of caps. By Proposition 2.16, the radii of these caps cannot tend to zero. By Corollary 2.18, such a sequence has a subsequence that converges in $L^2(\mathbb{S}^2)$. The limit of such a subsequence is an extremal. □

Not a Step. As explained above in Step 2, the fundamental potential obstruction to the precompactness of (nonnegative) extremizing sequences was the possibility that $|f_\nu|^2$ could converge weakly to a Dirac mass, or to a sum of two Dirac masses at a pair of antipodal points. Exclusion of this possibility relied on a suitable lower bound for **S** relative to **P**. The following result examines a natural one-parameter family of candidate trial functions, which provide an alternative source for a lower bound for **S**.

Proposition 2.19. *For all $\xi \in \mathbb{R}^3$ with $|\xi|$ sufficiently large,*

$$\|\widehat{e_\xi \sigma}\|_{L^4(\mathbb{R}^3)} > \mathcal{R}_{\mathbb{P}^2} \|e_\xi\|_{L^2(\mathbb{S}^2)}.$$

If $\xi = (0, 0, \lambda)$, then $e_\xi^2 / \|e_\xi\|_2^2$ does converge weakly as $\lambda \to +\infty$ to a constant multiple of a Dirac mass at $(0, 0, 1)$. Proposition 2.19 is proved in Section 17 via a perturbative calculation.

By taking the considerations of Step 2 involving even functions into account, Proposition 2.19 provides an alternative route to the essential comparison $\mathbf{S} > (3/2)^{1/4}\mathbf{P}$. Although Proposition 2.19 is not strictly necessary for the main lines of our proof, the calculation that underlies it is a natural tool for the investigation of manifolds more general than $\mathbb{S}^2$. However, both routes rely on specific properties of the sphere and paraboloid, whose generalization to related problems is not certain.

3. Step 2: $\mathbf{S} \geq (3/2)^{1/4}\mathbf{P}$

We begin by establishing the comparison $\mathbf{S} \geq \mathbf{P}$. This is based directly on the fact that a sphere is osculated to second order by an appropriate paraboloid.

Lemma 3.1. *The optimal constants $\mathbf{S}$ and $\mathbf{P}$, for $\mathbb{S}^2$ and $\mathbb{P}^2$ respectively, satisfy $\mathbf{S} \geq \mathbf{P}$. Moreover, for any $r, \varepsilon > 0$ and any $z \in \mathbb{S}^2$, there exists a function g supported in a cap $\mathcal{C}(z, r) \subset \mathbb{S}^2$ satisfying*

$$\|g\sigma * g\sigma\|_{L^2(\mathbb{R}^3)} \geq (\mathbf{P} - \varepsilon)^2 \|g\|_{L^2(\sigma)}^2,$$

where $L^2(\sigma)$ denotes $L^2(\mathbb{S}^2, \sigma)$.

Proof. Rotations are symmetries of the inequality (2-1). That is, for any rotation A of $\mathbb{R}^3$ and any $g \in L^2(\sigma)$, the function $g_A = g \circ A$ satisfies $\|g_A\|_{L^2(\sigma)} = \|g\|_{L^2(\sigma)}$ and $\|g_A\sigma * g_A\sigma\|_{L^2(\mathbb{R}^3)} = \|g\sigma * g\sigma\|_{L^2(\mathbb{R}^3)}$. Therefore it is no loss of generality to assume that $z = (0, 0, 1)$.

Write $x = (x', x_3) \in \mathbb{R}^2 \times \mathbb{R}$ as coordinates for $\mathbb{R}^3$. Each of the two convolution inequalities under consideration here (one for $\mathbb{S}^2$, one for $\mathbb{P}^2$) is equivalent to a corresponding adjoint Fourier restriction inequality, with optimal constants $\mathcal{R}$ and $\mathcal{R}_{\mathbb{P}^2}$ respectively. It suffices to prove that for each $\varepsilon > 0$, there exists f_ε supported in the set of all $(x', x_3) \in \mathbb{S}^2$ such that $|x'| < \varepsilon$ and $x_3 > 0$, such that $\|f_\varepsilon\|_{L^2(\sigma)} \leq 1+\varepsilon$ and $\|\widehat{f_\varepsilon \sigma}\|_{L^4(\mathbb{R}^3)} \geq (\mathcal{R}_{\mathbb{P}^2} - \varepsilon)$.

By definition of $\mathcal{R}_{\mathbb{P}^2}$, for any $\varepsilon > 0$ there exists a compactly supported C^∞ function $F_\varepsilon : \mathbb{R}^2 \to \mathbb{R}$ satisfying

$$\int_{\mathbb{R}^2} |F_\varepsilon(x_1, x_2)|^2 \, dx_1 \, dx_2 = 1 \quad \text{and} \quad \int_{\mathbb{R}^3} |\widehat{F_\varepsilon \sigma_P}|^4 \geq (\mathcal{R}_{\mathbb{P}^2} - \varepsilon)^4.$$

Here we have mildly abused notation in that the domain of F_ε is not $\mathbb{P}^2$; by $\widehat{F_\varepsilon \sigma_P}(y', y_3)$ we mean $\int_{\mathbb{R}^2} F_\varepsilon(x') e^{-ix'\cdot y'} e^{-iy_3|x'|^2/2} \, dx'$.

Suppose that F_ε is supported in $\{x' \in \mathbb{R}^2 : |x'| \leq \rho_\varepsilon\}$, where $\rho_\varepsilon \geq 1$. For $\delta \in (0, \varepsilon\rho_\varepsilon^{-1}]$ and $(x', x_3) \in \mathbb{S}^2$, define

$$f_{\varepsilon,\delta}(x', x_3) = \delta^{-1} F_\varepsilon(\delta^{-1} x').$$

Then $f_{\varepsilon,\delta}$ is supported in $\mathcal{C}((0, 0, 1), \delta\rho_\varepsilon) \subset \mathcal{C}((0, 0, 1), \varepsilon)$. Because $d\sigma(x) = (1 + O(\varepsilon^2))dx'$ in $\mathcal{C}((0, 0, 1), \varepsilon)$, we have $\|f_{\varepsilon,\delta}\|_{L^2(\sigma)} = (1 + O(\varepsilon))$.

Now

$$\begin{aligned}\widehat{f_{\varepsilon,\delta}\sigma}(y) &= \int_{\mathbb{R}^2} f_{\varepsilon,\delta}(x', \sqrt{1-|x'|^2}) e^{-iy'\cdot x'} e^{-iy_3\sqrt{1-|x'|^2}} h(x')\,dx' \\ &= \int_{\mathbb{R}^2} \delta^{-1} F_\varepsilon(\delta^{-1}x') e^{-iy'\cdot x'} e^{-iy_3\sqrt{1-|x'|^2}} h(x')\,dx' \\ &= \delta e^{-iy_3} \int_{\mathbb{R}^2} F_\varepsilon(x') e^{-i\delta y'\cdot x'} e^{-iy_3\left(\sqrt{1-\delta^2|x'|^2}-1\right)} h(\delta x')\,dx',\end{aligned}$$

where $h = d\sigma/dx'$ satisfies $h(x') = 1 + O(|x'|^2)$. Substitute $(u', u_3) = (\delta y', -\delta^2 y_3)$ and let $g_{\varepsilon,\delta}(u) = \delta^{-1} e^{iy_3}\widehat{f_{\varepsilon,\delta}\sigma}(y)$. Then $\|\widehat{f_{\varepsilon,\delta}\sigma}\|_{L^4(\mathbb{R}^3,\,dy)} = \|g_{\varepsilon,\delta}\|_{L^4(\mathbb{R}^3\,du)}$, and

$$g_{\varepsilon,\delta}(u) = \int_{\mathbb{R}^2} F_\varepsilon(x') e^{-iu'\cdot x'} e^{i\delta^{-2}u_3(\sqrt{1-\delta^2|x'|^2}-1)} h(\delta x')\,dx'.$$

Expanding as

$$\delta^{-2}\left(\sqrt{1-\delta^2|x'|^2}-1\right) = -\tfrac{1}{2}|x'|^2 + O(\delta^2|x'|^4)$$

gives

$$g_{\varepsilon,\delta}(u) = \int_{\mathbb{R}^2} F_\varepsilon(x') e^{-iu'\cdot x'} e^{-iu_3|x'|^2/2}\left(1 + O(\delta^2|x'|^2 + \delta^2|x'|^4)\right)dx'.$$

Let $\lambda < \infty$ be another parameter. Then uniformly for all u satisfying $|u| \le \lambda$,

$$g_{\varepsilon,\delta}(u) = \widehat{F_\varepsilon\sigma_P}(u) + O(\delta^2\rho_\varepsilon^4).$$

Therefore with ε, λ fixed,

$$\limsup_{\delta\to 0} \|\widehat{f_{\varepsilon,\delta}\sigma}\|_{L^4(\mathbb{R}^3)}^4 \ge \int_{|u|\le\lambda} |\widehat{F_\varepsilon\sigma_P}(u)|^4\,du.$$

Therefore

$$\limsup_{\delta\to 0} \|\widehat{f_{\varepsilon,\delta}\sigma}\|_{L^4(\mathbb{R}^3)} \ge \|\widehat{F_\varepsilon\sigma_P}\|_{L^4(\mathbb{R}^3)} \ge (\mathcal{R}_{\mathbb{P}^2} - \varepsilon),$$

while

$$\|f_{\varepsilon,\delta}\|_{L^2(\sigma)} = 1 + O(\varepsilon). \qquad \square$$

Improvement by the factor $(3/2)^{1/4}$ is based on the reflection symmetry $x \mapsto -x$ of $\mathbb{S}^2$. Recall $\tilde{f}(x) = \overline{f(-x)}$, which simplifies to $\tilde{f}(x) = f(-x)$ for real-valued functions. Denote by $\langle F, G\rangle$ the pairing of two functions in $L^2(\mathbb{R}^3)$, that is, $\langle F, G\rangle = \int_{\mathbb{R}^3} F\overline{G}dx$.

Lemma 3.2. *For any four real-valued functions $f_j \in L^2(\mathbb{S}^2)$,*

$$\langle f_1\sigma * f_2\sigma, f_3\sigma * f_4\sigma\rangle = \langle f_1\sigma * \tilde{f}_3\sigma, \tilde{f}_2\sigma * f_4\sigma\rangle \tag{3-1}$$

and

$$\|f_1\sigma * f_2\sigma\|_{L^2(\mathbb{R}^3)} = \|f_1\sigma * \tilde{f}_2\sigma\|_{L^2(\mathbb{R}^3)}. \tag{3-2}$$

Proof. The inequality $\|f\sigma * g\sigma\|_{L^2(\mathbb{R}^3)} \le \mathbf{S}^2\|f\|_{L^2(\sigma)}\|g\|_{L^2(\sigma)}$ ensures that these quantities are well defined, and that the first identity holds for all L^2 functions provided that it holds for all nonnegative continuous functions f_j. In that case $f_3\sigma * f_4\sigma(x) \le C|x|^{-1}$ for all $x \in \mathbb{R}^3$, where $C < \infty$ depends on f_3, f_4, and $f_3\sigma * f_4\sigma$ is continuous except at $x = 0$. For real-valued functions $F \in C^0(\mathbb{R}^3)$ and $f_j \in C^0(\mathbb{S}^2)$,

$$\langle f_1\sigma * f_2\sigma,\ F\rangle = \int (\tilde{f}_2\sigma * F) f_1\, d\sigma,$$

a consequence of the definition of convolution of measures and Fubini's theorem. Limiting arguments then lead to (3-1).

Equation (3-2) now follows:

$$\begin{aligned}\|f_1\sigma * f_2\sigma\|^2_{L^2(\mathbb{R}^3)} &= \langle f_1\sigma * f_2\sigma,\ f_1\sigma * f_2\sigma\rangle = \langle f_1\sigma * f_2\sigma,\ f_2\sigma * f_1\sigma\rangle \\ &= \langle f_1\sigma * \tilde{f}_2\sigma,\ \tilde{f}_2\sigma * f_1\sigma\rangle = \langle f_1\sigma * \tilde{f}_2\sigma,\ f_1\sigma * \tilde{f}_2\sigma\rangle = \|f_1\sigma * \tilde{f}_2\sigma\|^2_{L^2}. \qquad \square\end{aligned}$$

Proof of Lemma 2.2. Let $g \in L^2(\mathbb{S}^2)$ be supported in $\{x : x_3 > \frac{1}{2}\}$. Set $d\mu = g\, d\sigma$. Let $f(x) = 2^{-1/2}\big(g(x) + \overline{g(-x)}\big)$ and $d\nu = f\, d\sigma = 2^{-1/2}\big(\mu + \tilde{\mu}\big)$. The two terms $g(x)$ and $g(-x)$ have disjoint supports, so

$$\|f\|^2_{L^2(\mathbb{S}^2)} = \|g\|^2_{L^2(\mathbb{S}^2)}.$$

Now

$$\nu * \nu = \tfrac{1}{2}(\mu + \tilde{\mu}) * (\mu + \tilde{\mu}) = \tfrac{1}{2}((\mu * \mu) + (\tilde{\mu} * \tilde{\mu}) + 2(\mu * \tilde{\mu})).$$

The three summands on the right side have pairwise disjoint supports; the first is supported where $x_3 > 1$, the second where $x_3 < -1$, and the third where $|x_3| < 1$. Therefore

$$\|\nu * \nu\|^2_{L^2(\mathbb{R}^3)} = \tfrac{1}{4}(\|\mu * \mu\|^2_{L^2} + \|\tilde{\mu} * \tilde{\mu}\|^2_{L^2} + 4\|\mu * \tilde{\mu}\|^2_{L^2}).$$

There holds $\|\mu * \mu\|_{L^2} = \|\tilde{\mu} * \tilde{\mu}\|_{L^2}$, since one is the reflection about the origin of the other. By Lemma 3.2, it is also the case that $\|\mu * \tilde{\mu}\|^2_{L^2} = \|\mu * \mu\|^2_{L^2}$. Thus

$$\|\nu * \nu\|^2_{L^2(\mathbb{R}^3)} = \tfrac{3}{2}\|\mu * \mu\|^2_{L^2},$$

establishing Lemma 2.2. $\square$

Proof of Corollary 2.3. Let $\varepsilon > 0$. Choose $g \in L^2(\mathbb{S}^2)$, supported in $\{x \in \mathbb{S}^2 : x_3 > \frac{1}{2}\}$, satisfying $\|g\sigma * g\sigma\|_2^2 \ge (\mathbf{P} - \varepsilon)^4\|g\|^4_{L^2(\mathbb{S}^2)}$. By replacing g by $|g|$, we may assume that $g \ge 0$.

Consider once more $f(x) = 2^{-1/2}\big(g(x) + g(-x)\big)$. By Lemma 2.2,

$$\|f\sigma * f\sigma\|^2_{L^2(\mathbb{R}^3)} = \tfrac{3}{2}\|g\sigma * g\sigma\|^2_{L^2(\mathbb{R}^3)} \ge \tfrac{3}{2}(\mathbf{P} - \varepsilon)^4\|g\|^4_{L^2(\mathbb{S}^2)} = \tfrac{3}{2}(\mathbf{P} - \varepsilon)^4\|f\|^4_{L^2(\mathbb{S}^2)}.$$

Letting $\varepsilon \to 0$ yields Corollary 2.3. $\square$

4. Step 3: $\mathbf{S} \geq 2^{1/4}\mathbf{P}$

Proof of Lemma 2.4. We will obtain a lower bound for **S** by calculating $\|f\sigma * f\sigma\|_2^2$ for $f \equiv 1$. The following facts are well known: The unit ball in $\mathbb{R}^3$ has volume $4\pi/3$, $\sigma(\mathbb{S}^2) = 4\pi$, and the volume form in $\mathbb{R}^3$ in polar coordinates is $r^2\,dr\,d\sigma(\theta)$.

One calculates that

$$\sigma * \sigma(x) = a|x|^{-1}\chi_{|x|\leq 2} \tag{4-1}$$

for a certain constant $a > 0$. We will not need to evaluate a, which will cancel out at the end of the calculation. Let σ_P denote the measure dx' on the paraboloid $\mathbb{P}^2 = \{x \in \mathbb{R}^3 : x_3 = \frac{1}{2}|x'|^2\}$. What we do need to know is that

$$\sigma_P * \sigma_P(z) \equiv \tfrac{1}{2}a\chi_\Omega$$

where Ω denotes the support of $\sigma_P * \sigma_P$ and this constant a is the same as that in (4-1). This factor of $\frac{1}{2}$ in the definition of $\mathbb{P}^2$ is required to make the curvature of $\mathbb{P}^2$ equal to the curvature of $\mathbb{S}^2$; one sees that they are equal by writing the equation for $\mathbb{S}^2$ near the north pole as $x_3 - 1 = (1 - |x'|^2)^{1/2} - 1$ and Taylor expanding the right side. Note that the factor $a/2$ in the formula for $\sigma_P * \sigma_P$ agrees with the limit as $|x| \to 2$ of the function $a/|x|$, which appears in the formula for $\sigma * \sigma$. This asymptotic equality must hold since the two surfaces have equal curvatures; hence the two convolutions must agree on the diagonal of the maps $(x, y) \mapsto x + y$. We will not prove that $\sigma_P * \sigma_P$ is constant on its support; this is a reflection of the symmetry of the paraboloid (including appropriate dilation symmetry) and invariance of curvature under mappings of the form $(x', x_3) \mapsto (x', x_3 - L(x'))$ where $L : \mathbb{R}^2 \to \mathbb{R}^1$ is linear.

The support of $\sigma_P * \sigma_P$ is

$$\Omega = \{z : z_3 > \tfrac{1}{4}|z'|^2\}.$$

It is known [Foschi 2007; Hundertmark and Zharnitsky 2006] that any Gaussian is an extremizer for the paraboloid, and conversely. Another proof that Gaussians extremize the inequality is in [Bennett et al. 2009]. Set $F(x', x_3) = e^{-|x'|^2/2} \equiv e^{-x_3}$ on the paraboloid. Observe that if $x + y = z \in \mathbb{R}^3$, then

$$F(x)F(y) = e^{-x_3 - y_3} = e^{-z_3}.$$

Therefore

$$(F\sigma_P * F\sigma_P)(z) = \tfrac{1}{2}ae^{-z_3}\chi_{z_3 > |z'|^2/4}.$$

Consequently

$$\begin{aligned}\|F\sigma_P * F\sigma_P\|_2^2 &= \tfrac{1}{4}a^2 \int_{z' \in \mathbb{R}^2} \int_{z_3 > |z'|^2/4} e^{-2z_3}\,dz \\ &= \tfrac{1}{4}a^2 \int_0^\infty 2\pi \int_{r^2/4}^\infty e^{-2s}\,ds\,r\,dr = \tfrac{1}{4}a^2 2\pi \int_0^\infty \tfrac{1}{2}e^{-r^2/2}\,r\,dr = \tfrac{1}{4}\pi a^2.\end{aligned}$$

On the other hand,

$$\|\sigma * \sigma\|_{L^2(\mathbb{R}^3)}^2 = \int_{|x|\leq 2} a^2|x|^{-2}\,dx = a^2 \int_0^2 r^{-2}\,4\pi r^2\,dr = 4\pi a^2 \int_0^2 dr = 8\pi a^2.$$

Meanwhile

$$\|1\|^2_{L^2(\sigma)} = \sigma(\mathbb{S}^2) = 4\pi,$$

and

$$\|F\|^2_{L^2(\sigma_P)} = \int_{\mathbb{R}^2} e^{-2|x|^2/2}\,dx = \int_0^\infty e^{-r^2} 2\pi r\,dr = \pi.$$

Putting this all together,

$$\frac{\|F\sigma_P * F\sigma_P\|_2^2}{\|F\|^4_{L^2(\sigma_P)}} = \frac{a^2\pi/4}{\pi^2} = \frac{a^2}{4\pi},$$

while

$$\frac{\|1\sigma * 1\sigma\|_2^2}{\|1\|^4_{L^2(\sigma)}} = \frac{8\pi a^2}{(4\pi)^2} = \frac{a^2}{2\pi}.$$

The second ratio is equal to twice the first, as claimed. □

5. Step 4: Symmetrization

Proposition 2.7 states that $\|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)} \le \|f_\star\sigma * f_\star\sigma\|_{L^2(\mathbb{R}^3)}$ for any nonnegative function $f \in L^2(\mathbb{S}^2)$, where $f_\star$ denotes the antipodally symmetric rearrangement of f, defined in Definition 2.6.

Proof of Proposition 2.7. Let σ denote surface measure on $\mathbb{S}^2$. For $h \ge 0$,

$$\|h\sigma * h\sigma\|^2_{L^2} = \int h(a)h(b)h(c)h(d)\,d\lambda(a,b,c,d) \tag{5-1}$$

for a certain nonnegative measure λ that is supported on the set where $a+b=c+d$, and that is invariant under the transformations

$$(a,b,c,d) \mapsto (b,a,c,d), \qquad (a,b,c,d) \mapsto (a,-c,-b,d)$$
$$(a,b,c,d) \mapsto (c,d,a,b), \qquad (a,b,c,d) \mapsto (-a,-b,-c,-d).$$

This invariance, which is essential to the discussion, follows from the identities

$$f\sigma * g\sigma = g\sigma * f\sigma,$$
$$\langle f\sigma * g\sigma, h\sigma * k\sigma\rangle = \langle h\sigma * k\sigma, f\sigma * g\sigma\rangle,$$
$$\langle f\sigma * g\sigma, h\sigma * k\sigma\rangle = \langle f\sigma * \tilde{h}\sigma, \tilde{g}\sigma * k\sigma\rangle$$

for arbitrary real-valued functions, where $\tilde{F}(x) = F(-x)$.

Denote by G the finite group of symmetries of $(\mathbb{R}^3)^4$ that these generate. G has cardinality 48. Indeed, exactly one of a and $-a$ appears; suppose that a appears. There are 4 places in which it can go. Then $\pm b$ can go into any of 3 slots, but whether it is $+b$ or $-b$ is determined by which slot. There remain two slots into which $\pm c$ can go; again, the $\pm$ sign is determined by the slot. Then $\pm d$ goes into the remaining slot, with the $\pm$ sign again determined. The analysis is parallel if $-a$ appears. Thus there are $2\times 4\times 3\times 2 = 48$ possibilities.

By the orbit of a point we mean its image under G; by a generic point we mean one whose orbit has cardinality 48. In (5-1), it suffices to integrate only over all *generic* 4-tuples (a, b, c, d) satisfying $a+b=c+d$, since these form a set of full λ-measure.

To the orbit $\mathbb{O}$ we associate the functions

$$\mathcal{F}(\mathbb{O}) = \sum_{(a,b,c,d)\in\mathbb{O}} f(a)f(b)f(c)f(d) \quad \text{and} \quad \mathcal{F}_\star(\mathbb{O}) = \sum_{(a,b,c,d)\in\mathbb{O}} f_\star(a)f_\star(b)f_\star(c)f_\star(d).$$

Let Ω denote the set of all orbits of generic points. We can write

$$\|f*f\|_{L^2}^2 = \int_\Omega \mathcal{F}(\mathbb{O})\,d\tilde{\lambda}(\mathbb{O}) \quad \text{and} \quad \|f_\star * f_\star\|_{L^2}^2 = \int_\Omega \mathcal{F}_\star(\mathbb{O})\,d\tilde{\lambda}(\mathbb{O})$$

for a certain nonnegative measure $\tilde{\lambda}$. Therefore it suffices to prove that for any generic orbit $\mathbb{O}$,

$$\sum_{(a,b,c,d)\in\mathbb{O}} f(a)f(b)f(c)f(d) \le \sum_{(a,b,c,d)\in\mathbb{O}} f_\star(a)f_\star(b)f_\star(c)f_\star(d). \tag{5-2}$$

Fix any generic ordered 4-tuple (a, b, c, d) satisfying $a+b=c+d$. We prove (5-2) for its orbit. By homogeneity, it is no loss of generality to assume that $f^2(a)+f^2(-a)=1$ and that the same holds simultaneously for b, c, d. Thus we may write

$$f(a)=\cos(\varphi), \quad f(b)=\cos(\psi), \quad f(c)=\cos(\alpha), \quad f(d)=\cos(\beta)$$

for some $\varphi, \psi, \alpha, \beta \in [0, \pi/2]$ with $f(-a)=\sin(\varphi), \ldots, f(-d)=\sin(\beta)$. This means that

$$f_\star(x) = 2^{-1/2} \quad \text{for each } x \in \{\pm a, \pm b, \pm c, \pm d\}.$$

Now

$$\begin{aligned}\frac{1}{8}\sum_{(a',b',c',d')\in\mathbb{O}} f(a')f(b')f(c')f(d') &= \cos(\varphi)\cos(\psi)\cos(\alpha)\cos(\beta)+\sin(\varphi)\sin(\psi)\sin(\alpha)\sin(\beta)\\ &\quad +\cos(\varphi)\sin(\psi)\cos(\alpha)\sin(\beta)+\cos(\varphi)\sin(\psi)\sin(\alpha)\cos(\beta)\\ &\quad +\sin(\varphi)\cos(\psi)\cos(\alpha)\sin(\beta)+\sin(\varphi)\cos(\psi)\sin(\alpha)\cos(\beta)\\ &= \Gamma(\varphi,\psi,\alpha,\beta),\end{aligned}$$

where

$$\Gamma(\varphi,\psi,\alpha,\beta) = \cos(\varphi)\cos(\psi)\cos(\alpha)\cos(\beta)+\sin(\varphi)\sin(\psi)\sin(\alpha)\sin(\beta)+\sin(\varphi+\psi)\sin(\alpha+\beta).$$

Therefore the following lemma will complete the proof of Proposition 2.7. □

Lemma 5.1. $\max_{\varphi,\psi,\alpha,\beta\in[0,\pi/2]}\Gamma(\varphi,\psi,\alpha,\beta) = \frac{3}{2}$. *Moreover, this maximum value is attained only at* $(\frac{\pi}{4},\frac{\pi}{4},\frac{\pi}{4},\frac{\pi}{4})$.

Since

$$\Gamma(\tfrac{\pi}{4},\tfrac{\pi}{4},\tfrac{\pi}{4},\tfrac{\pi}{4}) = 1+(1/\sqrt{2})^4+(1/\sqrt{2})^4 = \tfrac{3}{2},$$

the maximum value of Γ is at least $\frac{3}{2}$. This point corresponds to the values taken by $f_\star$. Compare this with $\Gamma(0,0,0,0)=1$, which represents the extreme case when f vanishes at one of each pair of antipodal points; this ratio $(3/2)/1$ is the same $3/2$ that appears in Corollary 2.3.

Proof. We write Γ as

$$\begin{aligned}\Gamma &= \cos(\phi+\psi)\cos(\alpha+\beta)+\sin(\phi+\psi)\sin(\alpha+\beta)+\cos\phi\cos\psi\sin\alpha\sin\beta+\sin\phi\sin\psi\cos\alpha\cos\beta\\ &= \cos((\phi+\psi)-(\alpha+\beta))+\cos\phi\cos\psi\sin\alpha\sin\beta+\sin\phi\sin\psi\cos\alpha\cos\beta.\end{aligned}$$

Now

$$\begin{aligned}\cos\phi\cos\psi &= \frac{\cos(\phi+\psi)+\cos(\phi-\psi)}{2}\le\frac{1+\cos(\phi+\psi)}{2},\\ \sin\alpha\sin\beta &= \frac{-\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}\le\frac{1-\cos(\alpha+\beta)}{2}\end{aligned}$$

with equality only if $\phi=\psi$ and $\alpha=\beta$, and there are similar identities for $\sin\phi\sin\psi$ and $\cos\alpha\cos\beta$. Therefore

$$\begin{aligned}\Gamma &\le \cos((\phi+\psi)-(\alpha+\beta))+\tfrac{1}{4}(1+\cos(\phi+\psi))(1-\cos(\alpha+\beta))\\ &\qquad\qquad\qquad\qquad\qquad\qquad +\tfrac{1}{4}(1-\cos(\phi+\psi))(1+\cos(\alpha+\beta))\\ &= \cos((\phi+\psi)-(\alpha+\beta))+\tfrac{1}{2}(1-\cos(\phi+\psi)\cos(\alpha+\beta))\\ &= \cos((\phi+\psi)-(\alpha+\beta))-\tfrac{1}{2}\big(\cos((\phi+\psi)+(\alpha+\beta))+\cos((\phi+\psi)-(\alpha+\beta))\big)+\tfrac{1}{2}\\ &= \tfrac{1}{2}\big(\cos((\phi+\psi)-(\alpha+\beta))-\cos((\phi+\psi)+(\alpha+\beta))\big)+\tfrac{1}{2}\le\tfrac{3}{2}.\end{aligned}$$

The value $\frac{3}{2}$ can only be attained if all inequalities in this derivation are equalities. Equality in the final inequality forces $\phi+\psi+\alpha+\beta=\pi$ and $\phi+\psi=\alpha+\beta$. Together with the equalities $\phi=\psi$ and $\alpha=\beta$ already noted, these force $\phi=\psi=\alpha=\beta=\pi/4$. □

6. Step 5: Big pieces of caps

In this section we prove Lemma 2.9. While we are ultimately interested in establishing strong structural control of near-extremal functions, here we establish a weak connection between functions satisfying modest lower bounds $\|\widehat{f\sigma}\|_4\ge\delta\|f\|_2$, with $\delta>0$ arbitrarily small, and characteristic functions of caps.

For each integer $k\ge 0$ choose a maximal subset $\{z_k^j\}\subset\mathbb{S}^2$ satisfying $|z_k^j-z_k^i|\ge 2^{-k}$ for all $i\ne j$. Then for any $x\in\mathbb{S}^2$ there exists z_k^i such that $|x-z_k^i|\le 2^{-k}$; otherwise x could be adjoined to $\{z_k^j\}$, contradicting maximality. Therefore the caps $\mathscr{C}_k^j=\mathscr{C}(z_k^j,2^{-k+1})$ cover $\mathbb{S}^2$ for each k, and there exists $C<\infty$ such that for any k, no point of $\mathbb{S}^2$ belongs to more than C of the caps $\mathscr{C}_k^j$. The constant C is independent of k.

For $p\in[1,\infty)$, the X_p norm is defined by

$$\|f\|_{X_p}^4=\sum_{k=0}^{\infty}\sum_j 2^{-4k}\Big(|\mathscr{C}_k^j|^{-1}\int_{\mathscr{C}_k^j}|f|^p\Big)^{4/p}.$$

The factor 2^{-4k} can alternatively be written as $|\mathscr{C}_k^j|^2$.

Define also

$$\Lambda_{k,j}(f) = \Big(|\mathscr{C}_k^j|^{-1}\int_{\mathscr{C}_k^j}|f|\Big)\Big(|\mathscr{C}_k^j|^{-1}\int_{\mathbb{S}^2}|f|^2\Big)^{-1/2}.$$

By Hölder's inequality,

$$\Lambda_{k,j}(f) \le \Big(|\mathscr{C}_k^j|^{-1}\int_{\mathscr{C}_k^j}|f|^2\Big)^{1/2}\Big(|\mathscr{C}_k^j|^{-1}\int_{\mathbb{S}^2}|f|^2\Big)^{-1/2} = \|f\|_{L^2(\mathscr{C}_k^j)}/\|f\|_{L^2(\mathbb{S}^2)} \le 1.$$

It is shown in [Moyua et al. 1999, Lemma 4.4] that $L^2 \subset X_p$ for any $p < 2$. We will exploit the following refinement, which is very closely related to a result in Bégout and Vargas [2007], and whose somewhat tedious proof is deferred to Section 18.

Lemma 6.1. *For any $p \in [1, 2)$, there exist $C < \infty$ and $\gamma > 0$ such that for any $f \in L^2(\mathbb{S}^2)$,*

$$\|f\|_{X_p} \le C\|f\|_2 \sup_{k,j}(\Lambda_{k,j}(f))^{\gamma}.$$

Thus $\|f\|_{X_p} \le C_p\|f\|_2$ for any $f \in L^2(\mathbb{S}^2)$. Moreover, when the X_p norm is not significantly smaller than the L^2 norm, $\sup_{k,j}\Lambda_{k,j}(f)$ cannot be small.

Proposition 6.2 (Moyua, Vargas, and Vega [1999]). *There exist $C < \infty$ and $p \in (1, 2)$ such that for any $f \in L^2(\mathbb{S}^2)$,*

$$\|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \le C\|f\|_{X_p}.$$

This result contains Lemma 2.9 by an elementary argument, but we give the details for the sake of completeness.

Proof of Lemma 2.9. Let $\delta > 0$. Let $0 \ne f \in L^2(\mathbb{S}^2)$ and suppose that $\|\widehat{f\sigma}\|_{L^4(\mathbb{R}^3)} \ge \delta\|f\|_2$. For convenience, normalize so that $\|f\|_2 = 1$. The hypothesis, combined with the proposition and the lemma above, yields

$$\sup_{k,j}\Lambda_{k,j}(f) \ge c\delta^{1/\gamma}.$$

Fix k and j such that $\Lambda_{k,j}(f) \ge \frac{1}{2}c\delta^{1/\gamma}$. Henceforth write $\mathscr{C} = \mathscr{C}_k^j$. Thus

$$\int_{\mathscr{C}}|f| \ge c_0\delta^{1/\gamma}|\mathscr{C}|^{1/2},$$

where $c_0 > 0$ is a constant independent of f.

Let $R \ge 1$. Define $E = \{x \in \mathscr{C} : |f(x)| \le R\}$. Set $g = f\chi_E$ and $h = f - f\chi_E$. Then g and h have disjoint supports, $g + h = f$, g is supported on $\mathscr{C}$, and $\|g\|_\infty \le R$. Now $|h(x)| \ge R$ for almost every $x \in \mathscr{C}$ for which $h(x) \ne 0$, so

$$\int_{\mathscr{C}}|h| \le R^{-1}\int_{\mathscr{C}}|h|^2 \le R^{-1}\|f\|_2^2 = R^{-1}.$$

Define R by $R^{-1} = \frac{1}{2}c_0\delta^{1/\gamma}|\mathscr{C}|^{1/2}$. Then

$$\int_{\mathscr{C}}|g| = \int_{\mathscr{C}}|f| - \int_{\mathscr{C}}|h| \ge \tfrac{1}{2}c_0\delta^{1/\gamma}|\mathscr{C}|^{1/2}.$$

By Hölder's inequality, since g is supported on $\mathscr{C}$,

$$\|g\|_2 \geq |\mathscr{C}|^{-1+1/2}\|g\|_{L^1(\mathscr{C})} \geq c\delta^{1/\gamma} = c\delta^{1/\gamma}\|f\|_2.$$

Thus the decomposition $f = g + h$ satisfies the conclusions of Lemma 2.9, with η_δ proportional to $\delta^{1/\gamma}$, and C_δ proportional to $\delta^{-1/\gamma}$. □

7. Analytic preliminaries

On near-extremals.

Lemma 7.1. *Let $f = g + h \in L^2(\mathbb{S}^2)$. Suppose that $g \perp h$, $g \neq 0$, and that f is δ-nearly extremal for some $\delta \in (0, \frac{1}{4}]$. Then*

$$\frac{\|h\|_2}{\|f\|_2} \leq C \max\Bigl(\frac{\|h\sigma * h\sigma\|_2^{1/2}}{\|h\|_2}\ ,\ \delta^{1/2}\Bigr). \tag{7-1}$$

Here $C < \infty$ is a constant independent of g and h.

Proof. The inequality is invariant under multiplication of f by a positive constant, so we may assume without loss of generality that $\|g\|_2 = 1$. We may assume that $\|h\|_2 > 0$, since otherwise the conclusion is trivial. Define $y = \|h\|_2$ and

$$\eta = \|h\sigma * h\sigma\|_2^{1/2}/\mathbf{S}\|h\|_2.$$

If $\eta > \frac{1}{2}$, then (7-1) holds trivially with $C = 2/\mathbf{S}$, for the left side cannot exceed 1 since $f = g + h$ with $g \perp h$.

Since $\|f\sigma * f\sigma\|_2^{1/2}$ is a constant multiple of $\|\widehat{f\sigma}\|_4$, the functional $f \mapsto \|f\sigma * f\sigma\|_2^{1/2}$ satisfies the triangle inequality. Therefore

$$(1-\delta)^4\mathbf{S}^4\|f\|_2^4 \leq \|f\sigma * f\sigma\|_2^2 \leq \bigl(\|g\sigma * g\sigma\|_2^{1/2} + \|h\sigma * h\sigma\|_2^{1/2}\bigr)^4 \leq \mathbf{S}^4(1+\eta y)^4.$$

Since $g \perp h$, $\|f\|_2^2 = 1 + y^2$ and therefore

$$(1-\delta)(1+y^2)^{1/2} \leq 1 + \eta y.$$

Squaring gives

$$(1-2\delta)(1+y^2) \leq 1 + 2\eta y + \eta^2 y^2.$$

Since $\delta \in (0, \frac{1}{4}]$ and $\eta \leq \frac{1}{2}$,

$$\tfrac{1}{2}y^2 \leq 2\delta + 2\eta y + \eta^2 y^2 \leq 2\delta + 2\eta y + \tfrac{1}{4}y^2,$$

whence either $y^2 \leq 16\delta$ or $y \leq 16\eta$.

Substituting the definitions of y and η and majorizing $\|h\|_2/\|f\|_2$ by $\|h\|_2/\|g\|_2$ yields the stated conclusion. □

Simple bilinear convolution estimates.

Lemma 7.2. *Let $f \in L^2(\mathbb{S}^2)$ be nonnegative and satisfy $\|f\|_2 \le 1$. Let $z \in \mathbb{S}^2$ and $\varepsilon > 0$. Let $R \ge 1$ and $0 < \rho \le 1$. Then*

$$\|f\sigma * f\sigma\|_{L^2(\{|x|>2-\varepsilon\})} \le CR^2\varepsilon^{1/2}\rho + C\Big(\int_{f(x)\ge R} f^2(x)\,d\sigma(x)\Big)^{1/2} + C\Big(\int_{|x-z|\ge\rho} f^2(x)\,d\sigma(x)\Big)^{1/2}.$$

Proof. Decompose $f = g + h$ where g and h are nonnegative,

$$\|h\|_2 \le \Big(\int_{f(x)\ge R} f^2(x)\,d\sigma(x)\Big)^{1/2} + \Big(\int_{|x-z|\ge\rho} f^2(x)\,d\sigma(x)\Big)^{1/2}$$

and $\|g\|_2 \le 1$ and $\|g\|_\infty \le R$, and g is supported on $\{x \in \mathbb{S}^2 : |x - z| \le \rho\}$. Then

$$g\sigma * g\sigma(x) \le R^2\sigma * \sigma(x) \le CR^2|x|^{-1}$$

for $|x| < 2$, and equals 0 otherwise. Moreover, $g\sigma * g\sigma$ is supported in $\{x : |x - 2z| < 2\rho\}$. The $L^2(\mathbb{R}^3)$ norm of $|x|^{-1}1_{|x|\le 2}$ over the intersection of this region with $\{x : |x| > 2-\varepsilon\}$ is $\le C\rho\varepsilon^{1/2}$. This gives the bound $CR^2\rho\varepsilon^{1/2}$ for $\|g\sigma * g\sigma\|_2$. Since $\|g\|_2 \le 1$, the general inequality

$$\|F\sigma * G\sigma\|_{L^2(\mathbb{R}^3)} \le C\|F\|_2\|G\|_2$$

gives the required bound for both $g\sigma * h\sigma$ and $h\sigma * h\sigma$. □

Corollary 7.3. *Let $\{f_\nu\}$ be a sequence of real-valued functions that are upper even-normalized above with respect to a sequence of caps $\mathscr{C}_\nu$ of radii r_ν. If*

$$\delta_\nu / r_\nu^2 \to 0,$$

then

$$\int_{|x|>2-\delta_\nu} (|f_\nu|\sigma * |f_\nu|\sigma)^2\,dx \to 0 \quad \text{as } \nu \to \infty.$$

Lemma 7.4. *Let $f \in L^2(\mathbb{S}^2)$ be a function that is upper even-normalized with respect to a cap $\mathscr{C}$ of radius r. Then for all $R \ge 1$,*

$$\int_{R^{1/2}r\le|x|\le 2-Rr^2} |(f\sigma * f\sigma)(x)|^2\,dx \le \Psi(R),$$

where $\Psi(R) \to 0$ as $R \to \infty$, and Ψ depends only on the function Θ in the normalization inequalities (2-3) *and*(2-4), *not on r.*

Proof. It suffices to prove this for r small, R large, and Rr^2 uniformly bounded. Let $\mathscr{C} = \mathscr{C}(z, r)$ have center $z \in \mathbb{S}^2$. Let $A \in [1, \infty)$ and decompose $f = g_+ + h_+ + g_- + h_-$, where g_+, g_- are supported respectively in $\mathscr{C}(z, Ar)$ and $\mathscr{C}(-z, Ar)$, $\|h_+\|_2 \le \Theta(A)$ and $\|h_-\|_2 \le \Theta(A)$, where $\Theta(A) \to 0$ as $A \to \infty$.

Expand $f\sigma * f\sigma$ as a sum of the resulting 16 terms. The terms $g_+\sigma * g_+\sigma$ and $g_-\sigma * g_-\sigma$ are supported where $|x| > 2 - CA^2r^2$. If we choose A so that $CA^2 < R$, then these vanish identically in the region $|x| \le 2 - Rr^2$. The (two) terms $g_+\sigma * g_-\sigma$ are supported where $|x| \le CAr$. Therefore they also contribute nothing, provided that $CAr \le R^{1/2}r$.

Each of the remaining terms involves at least one factor of h_+ or of h_-. Since $\|F\sigma * G\sigma\|_{L^2(\mathbb{R}^3)} \le C\|F\|_2\|G\|_2$ for all $F, G \in L^2(\mathbb{S}^2)$, and since $g_\pm, h_\pm = O(1)$ in $L^2(\mathbb{S}^2)$ norm, each of these terms is $O(\|h_\pm\|_2)$. Therefore

$$\int_{R^{1/2}r \le |x| \le 2-Rr^2} |f\sigma * f\sigma(x)|^2\, dx \le C\Theta(A)^2$$

for any A that satisfies $CA^2 < R$. This completes the proof, provided that $Rr^2 = O(1)$. □

The set of all caps can be made into a metric space. Define the distance ρ from $\mathscr{C}(y, r)$ to $\mathscr{C}(y', r')$ to be the Euclidean distance from $(y/r, \log(1/r))$ to $(y'/r', \log(1/r'))$ in $\mathbb{R}^3 \times \mathbb{R}^+$. Note that for instance when $r = r'$, the distance is $r^{-1}|y - y'|$, so this distance has the natural scaling. If $y = y'$, then the distance is $|\log(r/r')|$; this has the natural property that it depends only on the *ratio* of the two radii. The definition ensures that this is truly a metric.

For any metric space (X, ρ) and any equivalence relation $\equiv$ on X, the function

$$\varrho([x], [y]) = \inf_{x' \in [x],\, y' \in [y]} \rho(x', y')$$

is a metric on the set of equivalence classes $X/\equiv$. Let $\mathcal{M}$ be the set of all caps $\mathscr{C} \subset \mathbb{S}^2$ modulo the equivalence relation $\mathscr{C} \equiv -\mathscr{C}$, where $-\mathscr{C} = \{-z : z \in \mathscr{C}\}$. Then the following defines a metric on $\mathcal{M}$.

Definition 7.5. For any two caps $\mathscr{C}, \mathscr{C}' \subset \mathbb{S}^2$,

$$\varrho([\mathscr{C}], [\mathscr{C}']) = \min(\rho(\mathscr{C}, \mathscr{C}'), \rho(-\mathscr{C}, \mathscr{C}')),$$

where $[\mathscr{C}]$ denotes the equivalence class $[\mathscr{C}] = \{\mathscr{C}, -\mathscr{C}\} \in \mathcal{M}$.

We will also write $\varrho(\mathscr{C}, \mathscr{C}') = \varrho([\mathscr{C}], [\mathscr{C}'])$.

Lemma 7.6. *For any $\varepsilon > 0$ there exists $\rho < \infty$ such that*

$$\|\chi_{\mathscr{C}}\sigma * \chi_{\mathscr{C}'}\sigma\|_{L^2(\mathbb{R}^3)} < \varepsilon|\mathscr{C}|^{1/2}|\mathscr{C}'|^{1/2}, \quad \textit{whenever } \varrho(\mathscr{C}, \mathscr{C}') > \rho.$$

Proof. Let $\mathscr{C} = \mathscr{C}(z, r)$ and $\mathscr{C}' = \mathscr{C}(z', r')$. Set $f = |\mathscr{C}|^{-1/2}\chi_{\mathscr{C}} \le Cr^{-1}\chi_{\mathscr{C}}$ and $f' = |\mathscr{C}'|^{-1/2}\chi_{\mathscr{C}'} \le Cr'^{-1}\chi_{\mathscr{C}'}$. Without loss of generality, $r' \le r$. We may suppose that $r' \ll 1$; otherwise the caps are not far apart. We will also assume at first that no points are nearly antipodal, that is, that $|x + x'| \ge \delta$ for all $x \in \mathscr{C}$ and $x' \in \mathscr{C}'$, for some fixed constant $\delta > 0$; we will return to this point later.

Consider first the case where $r \sim r'$. Then we may assume that $|z - z'| \ge 10r$, say. Then $f\sigma * f'\sigma$ has L^∞ norm $\le Cr^{-2} \cdot r/|z - z'|$, and is supported in a three-dimensional cylinder whose base has radius Cr and whose height is $\le Cr^2 + Cr|z - z'| \le Cr|z - z'|$. The volume of this cylinder is $\le Cr^3|z - z'|$. In all,

$$\|f\sigma * f'\sigma\|_{L^2(\mathbb{R}^3)} \le Cr^{-1}|z - z'|^{-1} \cdot r^{3/2}|z - z'|^{1/2} = C(r/|z - z'|)^{1/2},$$

which is small precisely when the caps are far apart.

Consider next the case where $r' \ll r$, and still $|z - z'| \geq 10r$. Then the L^∞ norm is no more than $Cr^{-1}r'^{-1} \cdot r'|z - z'|^{-1}$. The support is contained in a tubular neighborhood of a (translated) cap of radius Cr; this tubular neighborhood has width $\leq Cr'|z - z'|$. Hence the volume of the support is $\leq Cr^2r'|z - z'|$. Consequently

$$\|f\sigma * f'\sigma\|_{L^2(\mathbb{R}^3)} \leq Cr^{-1}r'^{-1}|z-z'|^{-1} \cdot rr'^{1/2}|z-z'|^{1/2} = C(r'/|z-z'|)^{1/2} \leq C(r'/r)^{1/2}.$$

Consider next the case where $r' \ll r$ and $|z - z'| \leq 10r$. It suffices to replace f by its restriction F to the complement of the cap $\mathscr{C}'^\star$ centered at z' of radius $10r^{3/4}r'^{1/4}$, since

$$\|f - F\|_2 \leq Cr^{-1}r^{3/4}r'^{1/4} = C(r'/r)^{1/4} \ll 1.$$

$F\sigma * f'\sigma$ is supported in a region of volume $\leq Cr^3r'$, and as is easily verified,

$$\|F\sigma * f'\sigma\|_\infty \leq Cr^{-1}r'^{-1} \cdot (r'/r^{3/4}r'^{1/4}) = Cr^{-7/4}r'^{-1/4}.$$

Therefore

$$\|F\sigma * f'\sigma\|_2 \leq Cr^{-7/4}r'^{-1/4} \cdot (r^3r')^{1/2} = Cr^{-1/4}r'^{1/4} \ll 1.$$

It only remains to handle caps that are nearly antipodal. But this follows from the nonantipodal case by the identity

$$\|f\sigma * g\sigma\|_2 = \|\tilde{f}\sigma * g\sigma\|_2, \quad \text{where } \tilde{f}(x) \equiv \overline{f(-x)}. \qquad \square$$

Fourier integral operators. Here we discuss another ingredient required for the proof of Lemma 12.2, certain estimates that rely on cancellation, in contrast to those in the preceding section.

For $0 < \rho \lesssim 1$, define $T_\rho : L^2(\mathbb{S}^2) \to L^2(\mathbb{S}^2)$ by

$$T_\rho f(x) = \int f(y)\, d\mu_{x,\rho}(y),$$

where $\mu_{x,\rho}$ is arc-length measure on the circle $\{y \in \mathbb{S}^2 : |y - x| = \rho\}$, normalized to be a probability measure.

Let Δ denote the spherical Laplacian.

Lemma 7.7. *We have*

$$\|T_\rho f\|_{L^2(\mathbb{S}^2)} \leq C\|(I - \rho^2\Delta)^{-1/4} f\|_{L^2(\mathbb{S}^2)} \tag{7-2}$$

uniformly for all $\rho > 0$ and all $f \in L^2(\mathbb{S}^2)$.

Sketch of proof. There are three elements in the proof of (7-2).

(i) Consider any fixed $\rho \in (0, 2)$. Define $\Phi_\rho(x, y) = |x - y|^2 - \rho^2$. Then the 3×3 matrix

$$\begin{pmatrix} 0 & \partial\Phi_\rho/\partial x \\ \partial\Phi_\rho/\partial y & \partial^2\Phi_\rho/\partial x\partial y \end{pmatrix} \tag{7-3}$$

is nonsingular for any (x, y) satisfying $\Phi_\rho(x, y) = 0$. This is a straightforward computation, easily done by taking advantage of rotational symmetry to reduce to a computation of Taylor expansions about $x = (0, 0, 1)$ and $y = (\cos(\theta), 0, \sin(\theta))$.

(ii) T_ρ is defined by integration against a smooth density on $\{(x, y) \in \mathbb{S}^2 \times \mathbb{S}^2 : \Phi_\rho(x, y) = 0\}$. As discussed on [Sogge 1993, pages 188–9], the nonsingularity of the matrix (7-3) implies that T_ρ is a Fourier integral operator of order $-(n-1)/2 = -1/2$ on $\mathbb{S}^n = \mathbb{S}^2$. Any such operator is smoothing of order $1/2$ in the scale of L^2 Sobolev spaces [Sogge 1993].

(iii) If T_ρ is rewritten with appropriate normalizations in coordinates adapted to any cap $\mathscr{C}(z, \rho)$, then the inequality holds uniformly in ρ. The only issue here is as $\rho \to 0$, but plainly in that situation there is a limiting operator on $\mathbb{R}^2$, $f \mapsto \int_{\mathbb{S}^1} f(x-y)\,d\mu(y)$, where μ is arc length measure on $\mathbb{S}^1 \subset \mathbb{R}^2$. This limiting operator is again a Fourier integral operator of order $-1/2$. It follows that the bounds are uniform after rescaling. Reversal of the rescaling introduces the factor ρ^2 to Δ in the inequality. □

The operators T_ρ are related to our bilinear convolutions: For $f \in L^2(\mathbb{S}^2)$ and $x \in \mathbb{R}^3$ satisfying $0 < |x| < 2$,

$$(f\sigma * \sigma)(x) = c|x|^{-1} T_\rho f(x/|x|),$$

where $\rho^2 + |x/2|^2 = 1$. Define $e_\xi(x) = e^{x\cdot\xi}$, for $x \in \mathbb{R}^3$ and $\xi \in \mathbb{C}$ (and in particular for $x \in \mathbb{S}^2$). There is the more general identity

$$(f\sigma * e_{i\xi}\sigma)(x) = e_{i\xi}(x)(e_{-i\xi} f\sigma * \sigma)(x) = c|x|^{-1} e_{i\xi}(x) T_\rho(e_{-i\xi} f)(x). \tag{7-4}$$

Suppose that $g \in L^2(\mathbb{S}^2)$ takes the form $g(x) = \int_H a(\xi) e_{i\xi}(x)\,d\nu(\xi)$, where $H \subset \mathbb{R}^3$ is a two-dimensional subspace, ν is Lebesgue measure on H, and $a \in L^2(H)$. Then

$$(f\sigma * g\sigma)(x) = c|x|^{-1} \int_H a(\xi) e_{i\xi}(x) T_\rho(e_{-i\xi} f)(x)\,d\nu(\xi).$$

For $t \in (0, 2)$, define $\rho(t) > 0$ by

$$\rho(t)^2 + (t/2)^2 = 1.$$

Then for any interval $I \subset (0, 2)$,

$$\begin{aligned} \int_{|x|\in I} |(f\sigma * g\sigma)(x)|^2\,dx &\le C \int_I t^{-2} \left\| \int_H |a(\zeta)| \cdot |T_{\rho(t)}(e_{-i\zeta} f)|\,d\zeta \right\|_{L^2(\mathbb{S}^2)}^2 t^2\,dt \\ &= C \int_I \left\| \int_H |a(\zeta)| \cdot |T_{\rho(t)}(e_{-i\zeta} f)|\,d\zeta \right\|_{L^2(\mathbb{S}^2)}^2 dt. \end{aligned} \tag{7-5}$$

Fourier coefficient estimates in terms of the spherical Laplacian. The following routine lemma is convenient because it provides an intrinsic characterization of expressions that arise in the analysis. The proof relies on the machinery of pseudodifferential operators, and is left to the reader.

Lemma 7.8. *Let $\mathscr{C}$ be a cap of radius $\varrho \le \frac{1}{2}$. Let ϕ be the rescaling map associated with $\mathscr{C}$. Let f be supported in $\mathscr{C} \cup (-\mathscr{C})$. Then for any $t \in \mathbb{R}$ and $0 < r \le \varrho$,*

$$C^{-1} \|(I - r^2\Delta)^{t/2} f\|_{L^2(\mathbb{S}^2)}^2 \le \int_{\mathbb{R}^2} |\widehat{\phi^* f}(\xi)|^2 (1 + |r\varrho^{-1}\xi|^2)^t\,d\xi \le C \|(I - r^2\Delta)^{t/2} f\|_{L^2(\mathbb{S}^2)}^2.$$

Here $C \in (0, \infty)$ depends on t but not on $f, r, \varrho, \mathscr{C}$.

8. Step 6A: A decomposition algorithm

The following iterative procedure may be applied to any nonnegative function $f \in L^2(\mathbb{S}^2)$ of positive norm.

Decomposition algorithm. Initialize by setting $G_0 = f$, and $\varepsilon_0 = 1/2$.

Step ν. The inputs for Step ν are a nonnegative function $G_\nu \in L^2(\mathbb{S}^2)$ and a positive number ε_ν. Its outputs are functions f_ν and $G_{\nu+1}$ and nonnegative numbers $\varepsilon_\nu^\star$ and $\varepsilon_{\nu+1}$. If $\|G_\nu\sigma * G_\nu\sigma\|_2 = 0$, then $G_\nu = 0$ almost everywhere. The algorithm then terminates, and we define $\varepsilon_\nu^\star = 0$ and $f_\nu = 0$, and $G_\mu = f_\mu = 0$ and $\varepsilon_\mu = 0$ for all $\mu > \nu$.

If $0 < \|G_\nu\sigma * G_\nu\sigma\|_2 < \varepsilon_\nu^2 \mathbf{S}^2 \|f\|_2^2$, then replace ε_ν by $\varepsilon_\nu/2$, and repeat until the first time that $\|G_\nu\sigma * G_\nu\sigma\|_2 \geq \varepsilon_\nu^2 \mathbf{S}^2 \|f\|_2^2$. Define $\varepsilon_\nu^\star$ to be this value of ε_ν. Then

$$(\varepsilon_\nu^\star)^2 \mathbf{S}^2 \|f\|_2^2 \leq \|G_\nu\sigma * G_\nu\sigma\|_2 \leq 4(\varepsilon_\nu^\star)^2 \mathbf{S}^2 \|f\|_2^2.$$

Apply Lemma 2.9 to obtain a cap $\mathscr{C}_\nu$ and a decomposition $G_\nu = f_\nu + G_{\nu+1}$ with disjointly supported nonnegative summands satisfying $f_\nu \leq C_\nu \|f\|_2 |\mathscr{C}_\nu|^{-1/2} \chi_{\mathscr{C}_\nu}$, and $\|f_\nu\|_2 \geq \eta_\nu \|f\|_2$. Here C_ν, η_ν are bounded above and below, respectively, by quantities that depend only on $\|G_\nu\sigma * G_\nu\sigma\|_2^{1/2} / \|G_\nu\|_2 \geq \varepsilon_\nu^\star$. Define $\varepsilon_{\nu+1} = \varepsilon_\nu^\star$, and move on to step $\nu + 1$. □

It is important for our application to observe that if f is even then at every step, f_ν may likewise be chosen to be even. The upper bound for f_ν then becomes

$$f_\nu \leq C_\nu |\mathscr{C}_\nu|^{-1/2} \chi_{\mathscr{C}_\nu \cup -\mathscr{C}_\nu}.$$

Henceforth the algorithm will be applied only to even functions, and we will always choose all f_ν to be even.

If the algorithm terminates at some finite step ν, then a finite decomposition $f = \sum_{k=0}^{\nu} f_k$ results.

Lemma 8.1. *Let $f \in L^2(\mathbb{S}^2)$ be a nonnegative function with positive norm. If the decomposition algorithm never terminates for f, then $\varepsilon_\nu^\star \to 0$ as $\nu \to \infty$, and $\sum_{\nu=0}^{N} f_\nu \to f$ in L^2 as $N \to \infty$.*

Proof. Assume without loss of generality that $\|f\|_2 = 1$. The functions f_ν have disjoint supports and hence are pairwise orthogonal, and $\sum_\nu f_\nu \leq f$, so $\sum_\nu \|f_\nu\|_2^2 \leq \|f\|_2^2$. Since the sequence $\varepsilon_\nu^\star$ is nonincreasing and $\|f_\nu\|_2 / \|f\|_2$ is bounded below by a function of $\varepsilon_\nu^\star$, this forces $\varepsilon_\nu^\star \to 0$.

The second conclusion is equivalent to $\|G_N\|_2 \to 0$. According to Lemma 2.9, $\|f_\nu\|_2$ is bounded below by a function of $\|G_\nu\sigma * G_\nu\sigma\|_2$. Since $\sum_\nu \|f_\nu\|_2^2 < \infty$, we have $\|f_\nu\|_2 \to 0$ and therefore $\|G_\nu\sigma * G_\nu\sigma\|_2 \to 0$. By construction, $G_{\nu+1}(x) \leq G_\nu(x)$ for every $x \in \mathbb{S}^2$, so $G(x) = \lim_{\nu\to\infty} G_\nu(x)$ exists and $\|G\sigma * G\sigma\|_2 \leq \|G_\nu\sigma * G_\nu\sigma\|_2$ for all ν. Thus $G\sigma * G\sigma \equiv 0$, so $G \equiv 0$. This forces $\|G_\nu\|_2 \to 0$, by the dominated convergence theorem. □

For general f, this decomposition may be highly inefficient. But if f is nearly extremal for the inequality (2-1), then more useful properties hold.

Lemma 8.2. *There exists a continuous function $\theta : (0,1] \to (0,\infty)$ such that for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any δ-nearly extremal nonnegative function $f \in L^2(\mathbb{S}^2)$ satisfying $\|f\|_2 = 1$, the functions f_ν and G_ν associated by the decomposition algorithm to f satisfy*

$$\|f_\nu\|_2 \ge \theta(\|G_\nu\|_2) \quad \textit{for any index } \nu \textit{ such that } \|G_\nu\|_2 \ge \varepsilon.$$

This is a direct consequence of Lemmas 2.9 and 7.1. It is essential for applications below that θ be independent of ε.

If f is nearly extremal, then the norms of f_ν and G_ν enjoy upper bounds independent of f, for all except very large ν.

Lemma 8.3. *There exist a sequence of positive constants $\gamma_\nu \to 0$ and a function $N : (0, \frac{1}{2}] \to \mathbb{Z}^+$ satisfying $N(\delta) \to \infty$ as $\delta \to 0$ such that for any nonnegative $f \in L^2(\mathbb{S}^2)$, if f is δ-nearly extremal then the quantities $\varepsilon_\nu^\star$ obtained when the decomposition algorithm is applied to f satisfy*

$$\begin{aligned} \varepsilon_\nu^\star &\le \gamma_\nu && \textit{for all } \nu \le N(\delta), \\ \|G_\nu\|_2 &\le \gamma_\nu \|f\|_2 && \textit{for all } \nu \le N(\delta), \\ \|f_\nu\|_2 &\le \gamma_\nu \|f\|_2 && \textit{for all } \nu \le N(\delta). \end{aligned}$$

This holds whether or not the algorithm terminates for f.

Proof.
$$\mathbf{S}^2\|G_\nu\|_2^2 \ge \|G_\nu\sigma * G_\nu\sigma\|_2 \ge \varepsilon_\nu^{\star 2}\mathbf{S}^2\|f\|_2^2 = (\varepsilon_\nu^{\star 2}\|f\|_2^2/\|G_\nu\|_2^2)\mathbf{S}^2\|G_\nu\|_2^2,$$

so $\varepsilon_\nu^\star \le \|G_\nu\|_2/\|f\|_2$. Thus the second conclusion implies the first. Since $\|f_\nu\|_2 \le \|G_\nu\|_2$, it also implies the third.

We recall two facts. First, Lemma 7.1, applied to $h = G_\nu$ and $g = f_0 + \cdots + f_{\nu-1}$, asserts that there are constants $c_0, C_1 \in \mathbb{R}^+$ such that if $f \in L^2$ is δ-nearly extremal, either $\|G_\nu\sigma * G_\nu\sigma\|_2 \ge c_0\|G_\nu\|_2^4\|f\|_2^{-2}$ or $\|G_\nu\|_2 \le C_1\delta^{1/2}\|f\|_2$. Second, according to Lemma 2.9, there exists a nondecreasing function $\rho : (0,\infty) \to (0,\infty)$ satisfying $\rho(t) \to 0$ as $t \to 0$ such that for every nonzero $f \in L^2$ and any ν, if $\|G_\nu\sigma * G_\nu\sigma\|_2 \ge t\|G_\nu\|_2^2$, then $\|f_\nu\|_2^2 \ge \rho(t)\|G_\nu\|_2^2$.

Choose a sequence $\{\gamma_\nu\}$ of positive numbers that tends monotonically to zero, but does so sufficiently slowly to satisfy

$$\nu\gamma_\nu^2\rho(c_0\gamma_\nu^2) > 1 \quad \text{for all } \nu.$$

Define $N(\delta)$ to be the largest integer satisfying $\gamma_{N(\delta)} \ge C_1\delta^{1/2}$. This $N(\delta) \to \infty$ as $\delta \to 0$ because $\gamma_\nu > 0$ for all ν.

Let f and δ be given. Suppose that $\nu \le N(\delta)$. We argue by contradiction, supposing that $\|G_\nu\|_2 > \gamma_\nu\|f\|_2$. Then $\|G_\nu\|_2 > C_1\delta^{1/2}\|f\|_2$ by definition of $N(\delta)$. By the dichotomy above,

$$\|G_\nu\sigma * G_\nu\sigma\|_2 \ge c_0\|G_\nu\|_2^4\|f\|_2^{-2} \ge c_0\gamma_\nu^2\|G_\nu\|_2^2.$$

By the second fact reviewed above,

$$\|f_\nu\|_2^2 \ge \rho(c_0\gamma_\nu^2)\|G_\nu\|_2^2 \ge \gamma_\nu^2\rho(c_0\gamma_\nu^2)\|f\|_2^2.$$

Since $\|G_\mu\|_2 \ge \|G_\nu\|_2$ for all $\mu \le \nu$, the same lower bound follows for $\|f_\nu\|_2^2$ for all $\mu \le \nu$. Since the functions f_μ are pairwise orthogonal, $\sum_{\mu\le\nu}\|f_\mu\|_2^2 \le \|f\|_2^2$, and consequently $\nu\gamma_\nu^2\rho(c_0\gamma_\nu^2) \le 1$, a contradiction. □

The next lemma also follows directly from the decomposition algorithm coupled with Lemma 2.9.

Lemma 8.4. *For any $\varepsilon > 0$ there exist $\delta_\varepsilon > 0$ and $C_\varepsilon < \infty$ such that for every δ_ε-nearly extremal nonnegative function $f \in L^2$, the functions f_ν and G_ν associated to f by the decomposition algorithm satisfy*

(i) *For any ν, if $\|G_\nu\|_2 \ge \varepsilon\|f\|_2$, then there exists a cap $\mathscr{C}_\nu \subset \mathbb{S}^2$ such that*

$$f_\nu \le C_\varepsilon \|f\|_2 |\mathscr{C}_\nu|^{-1/2} \chi_{\mathscr{C}_\nu \cup -\mathscr{C}_\nu}.$$

(ii) *If $\|G_\nu\|_2 \ge \varepsilon\|f\|_2$, then $\|f_\nu\|_2 \ge \delta_\varepsilon\|f\|_2$.*

9. Step 6B: A geometric property of the decomposition

We have established inequalities concerning the L^2 norms of the functions f_ν and G_ν that the decomposition algorithm yields, based on quite general principles and a single analytic fact, Lemma 2.9, concerning the particular inequality that we are studying. We next establish an additional inequality of a geometric nature, based on a single additional fact, the weak interaction of distant caps in the sense of Lemma 7.6.

Lemma 9.1. *In any metric space, for any N and r, any finite set S of cardinality N and diameter equal to r may be partitioned into two disjoint nonempty subsets $S = S' \cup S''$ such that* distance$(S', S'') \ge r/2N$. *Moreover, given two points $s', s'' \in S$ satisfying* distance$(s', s'') = r$, *this partition can be constructed so that $s' \in S'$ and $s'' \in S''$.*

Proof. Consider the metric balls B_k centered at s' of radii $kr/2N$ for $k = 1, 2, \dots, 2N$. By the pigeonhole principle, there exists k such that $(B_{k+1} \setminus B_k) \cap S = \varnothing$. Set $S' = B_k \cap S$ and $S'' = S \setminus S'$. The triangle inequality yields the conclusion. □

Lemma 9.2. *For any $\varepsilon > 0$ there exist $\delta > 0$ and $\lambda < \infty$ such that for any $0 \le f \in L^2(\mathbb{S}^2)$ that is δ-nearly extremal, the summands f_ν produced by the decomposition algorithm and the associated caps $\mathscr{C}_\nu$ satisfy*

$$\varrho(\mathscr{C}_j, \mathscr{C}_k) \le \lambda \quad \textit{whenever } \|f_j\|_2 \ge \varepsilon\|f\|_2 \textit{ and } \|f_k\|_2 \ge \varepsilon\|f\|_2.$$

Here ϱ is the distance between $\mathscr{C}_j \cup -\mathscr{C}_j$ and $\mathscr{C}_k \cup -\mathscr{C}_k$, as defined in Definition 7.5.

Proof. It suffices to prove this for all sufficiently small ε. Let f be a nonnegative L^2 function that satisfies $\|f\|_2 = 1$ and is δ-nearly extremal for a sufficiently small $\delta = \delta(\varepsilon)$, and let $\{G_\nu, f_\nu\}$ be associated to f via the decomposition algorithm. Set $F = \sum_{\nu=0}^N f_\nu$.

Suppose that $\|f_{j_0}\|_2 \ge \varepsilon$ and $\|f_{k_0}\|_2 \ge \varepsilon$. Let N be the smallest integer such that $\|G_{N+1}\|_2 < \varepsilon^3$. Since $\|G_\nu\|_2$ is a nonincreasing function of ν, and since $\|f_\nu\|_2 \le \|G_\nu\|_2$, necessarily $j_0, k_0 \le N$. Moreover, by Lemma 8.3, there exists $M_\varepsilon < \infty$ depending only on ε such that $N \le M_\varepsilon$. By Lemma 8.4, if δ is chosen to be a sufficiently small function of ε, then since $\|G_\nu\|_2 \ge \varepsilon^3$ for all $\nu \le N$, we have $f_\nu \le \theta(\varepsilon)|\mathscr{C}|^{-1/2}\chi_{\mathscr{C}\cup-\mathscr{C}}$ for all such ν, where θ is a continuous, strictly positive function on $(0, 1]$.

Now let $\lambda < \infty$ be a large quantity to be specified. It suffices to show that if $\delta(\varepsilon)$ is sufficiently small, an assumption that $\varrho(\mathscr{C}_j, \mathscr{C}_k) > \lambda$ implies an upper bound, which depends only on ε, for λ.

Lemma 9.1 yields a decomposition $F = F_1 + F_2 = \sum_{\nu \in S_1} f_\nu + \sum_{\nu \in S_2} f_\nu$, where $[0, N] = S_1 \cup S_2$ is a partition of $[0, N]$, $j_0 \in S_1$, $k_0 \in S_2$, and $\varrho(\mathscr{C}_j, \mathscr{C}_k) \geq \lambda/2N \geq \lambda/2M_\varepsilon$ for all $j \in S_1$ and $k \in S_2$. Certainly $\|F_1\|_2 \geq \|f_{j_0}\|_2 \geq \varepsilon$ and similarly $\|F_2\|_2 \geq \varepsilon$. The convolution cross term satisfies

$$\|F_1\sigma * F_2\sigma\|_2 \leq \sum_{j \in S_1} \sum_{k \in S_2} \|f_j\sigma * f_k\sigma\|_2 \leq M_\varepsilon^2 \gamma(\lambda/2M_\varepsilon)\theta(\varepsilon)^2,$$

where $\gamma(\lambda) \to 0$ as $\lambda \to \infty$ by Lemma 7.6. Therefore

$$\begin{aligned} \|F\sigma * F\sigma\|_2^2 &\leq \|F_1\sigma * F_1\sigma\|_2^2 + \|F_2\sigma * F_2\sigma\|_2^2 + C\|f\|_2^2 \|F_1\sigma * F_2\sigma\|_2 \\ &\leq \mathbf{S}^4 \|F_1\|_2^4 + \mathbf{S}^4 \|F_2\|_2^4 + M_\varepsilon^2 \gamma(\lambda/2M_\varepsilon)\theta(\varepsilon)^2. \end{aligned}$$

Since F_1 and F_2 have disjoint supports, $\|F_1\|_2^2 + \|F_2\|_2^2 \leq \|f\|_2^2 = 1$ and consequently

$$\|F_1\|_2^4 + \|F_2\|_2^4 \leq \max\left(\|F_1\|_2^2, \|F_2\|_2^2\right) \cdot \left(\|F_1\|_2^2 + \|F_2\|_2^2\right) \leq (1 - \varepsilon^2) \cdot 1 \leq 1 - \varepsilon^2.$$

Thus

$$\|F\sigma * F\sigma\|_2^2 \leq \mathbf{S}^4(1 - \varepsilon^2) + M_\varepsilon^2 \gamma(\lambda/2M_\varepsilon)\theta(\varepsilon)^2.$$

Therefore

$$\begin{aligned} (1-\delta)^2 \mathbf{S}^2 \leq \|f\sigma * f\sigma\|_2 &\leq \|F\sigma * F\sigma\|_2 + C\|f\|_2 \|f - F\|_2 \\ &\leq \|F\sigma * F\sigma\|_2 + C\varepsilon^3, \end{aligned}$$

so by transitivity

$$(1-\delta)^4 \mathbf{S}^4 \leq C\varepsilon^3 + \mathbf{S}^4(1 - \varepsilon^2) + M_\varepsilon^2 \gamma(\lambda/2M_\varepsilon)\theta(\varepsilon)^2.$$

Since $\gamma(t) \to 0$ as $t \to \infty$, for all sufficiently small $\varepsilon > 0$ this implies an upper bound, which depends only on ε, for λ, as was to be proved. □

10. Step 6C: Upper bounds for extremizing sequences

Proposition 2.14 states that any nearly extremal function satisfies appropriately scaled upper bounds relative to some cap. It is convenient for the proof to first observe that a superficially weaker statement implies the version stated.

Lemma 10.1. *There exists a function $\Theta : [1, \infty) \to (0, \infty)$ satisfying $\Theta(R) \to 0$ as $R \to \infty$ with the following property. For any $\varepsilon > 0$ and $\bar{R} \in [1, \infty)$ there exists $\delta > 0$ such that any nonnegative even function f that has $\|f\|_2 = 1$ and is δ-nearly extremal may be decomposed as $f = F + G$, where F and G are even and nonnegative with disjoint supports, $\|G\|_2 < \varepsilon$, and there exists a cap $\mathscr{C} = \mathscr{C}(z, r)$*

such that for any $R \in [1, \bar{R}]$,

$$\int_{\min(|x-z|,|x+z|)\geq Rr} F^2(x)\,d\sigma(x) \leq \Theta(R), \tag{10-1}$$

$$\int_{F(x)\geq Rr^{-1}} F^2(x)\,d\sigma(x) \leq \Theta(R). \tag{10-2}$$

Proof that Lemma 10.1 implies Proposition 2.14. Let Θ be the function promised by the lemma. Let ε and f be given, and assume without loss of generality that ε is small. Assuming as we may that Θ is a continuous, strictly decreasing function, define $\bar{R} = \bar{R}(\varepsilon)$ by the equation $\Theta(\bar{R}) = \varepsilon^2/2$. Let $\mathscr{C} = \mathscr{C}(z, r)$ and suppose $\delta = \delta(\varepsilon, \bar{R}(\varepsilon))$ along with F and G satisfy the conclusions of the lemma relative to ε and $\bar{R}(\varepsilon)$. Define χ to be the characteristic function of the set of all $x \in \mathbb{S}^2$ that satisfy either $\min(|x-z|, |x+z|) \geq \bar{R}r$ or $F(x) > \bar{R}|\mathscr{C}|^{-1/2}$. Redecompose $f = \tilde{F} + \tilde{G}$, where $\tilde{F} = (1-\chi)F$ and $\tilde{G} = G + \chi F$. Then $\|\tilde{G}\|_2 < 2\varepsilon$, while $\tilde{F}$ satisfies the required inequalities. For instance, if $R \leq \bar{R}$ then

$$\int_{\tilde{F}(x)\geq R|\mathscr{C}|^{-1/2}} \tilde{F}(x)^2\,d\sigma(x) \leq \int_{F(x)\geq R|\mathscr{C}|^{-1/2}} F(x)^2\,d\sigma(x) \leq \Theta(R),$$

while the integrand vanishes if $R > \bar{R}$. □

Proof of Lemma 10.1. Let $\eta : [1, \infty) \to (0, \infty)$ be a function to be chosen below, satisfying $\eta(t) \to 0$ as $t \to \infty$. This function will not depend on the quantity $\bar{R}$.

Let $\bar{R} \geq 1$, $R \in [1, \bar{R}]$, and $\varepsilon > 0$ be given. Let $\delta = \delta(\varepsilon, \bar{R}) > 0$ be a small quantity to be chosen below. Let $0 \leq f \in L^2(\mathbb{S}^2)$ be even and δ-nearly extremal. It is no loss of generality in normalizing such that $\|f\|_2 = 1$.

Let $\{f_\nu\}$ be the sequence of functions obtained by applying the decomposition algorithm to f. Choose $\delta = \delta(\varepsilon) > 0$ sufficiently small and $M = M(\varepsilon)$ sufficiently large to guarantee that $\|G_{M+1}\|_2 < \varepsilon/2$ and that f_ν and G_ν satisfy all conclusions of Lemma 8.4 and Lemma 8.3 for $\nu \leq M$. Set $F = \sum_{\nu=0}^{M} f_\nu$. Then $\|f - F\|_2 = \|G_{M+1}\|_2 < \varepsilon/2$.

Let $N \in \{0, 1, 2, \dots\}$ be the minimum of M and the smallest number such that $\|f_{N+1}\|_2 < \eta$. N is majorized by a quantity that depends only on η. Set $\mathscr{F} = \mathscr{F}_N = \sum_{k=0}^{N} f_k$. It follows from Lemma 8.4(ii) that

$$\|F - \mathscr{F}\|_2 < \gamma(\eta), \quad \text{where } \gamma(\eta) \to 0 \text{ as } \eta \to 0. \tag{10-3}$$

This function γ is independent of ε and $\bar{R}$.

To prove the lemma, we must produce an appropriate cap $\mathscr{C} = \mathscr{C}(z, r)$, and must establish the existence of Θ. To do the former is simple: To f_0 is associated a cap $\mathscr{C}_0 = \mathscr{C}(z_0, r_0)$ such that $f_0 \leq C|\mathscr{C}_0|^{-1/2}(\chi_{\mathscr{C}_0 \cup -\mathscr{C}_0})$. Then $\mathscr{C} = \mathscr{C}_0$ is the required cap. Note that by Lemma 2.9, $\|f_0\|_2 \geq c$ for some positive universal constant c.

Suppose that functions $R \mapsto \eta(R)$ and $R \mapsto \Theta(R)$ are chosen so that

$$\eta(R) \to 0 \text{ as } R \to \infty \quad \text{and} \quad \gamma(\eta(R)) \leq \Theta(R) \text{ for all } R.$$

Then by (10-3), $F - \mathcal{F}$ already satisfies the desired inequalities in $L^2(\mathbb{S}^2)$, so it suffices to show that $\mathcal{F}(x) \equiv 0$ whenever $\min(|x-z|, |x+z|) > Rr_0$, and that $\|\mathcal{F}\|_\infty \le R|\mathscr{C}_0|^{-1/2}$.

Each summand satisfies $f_k \le C(\eta)|\mathscr{C}_k|^{-1/2}\chi_{\mathscr{C}_k \cup -\mathscr{C}_k}$, where $C(\eta) < \infty$ depends only on η, and in particular, f_k is supported in $\mathscr{C}_k \cup -\mathscr{C}_k$. Now $\|f_k\|_2 \ge \eta$ for all $k \le N$ by definition of N. Therefore by Lemma 9.2, there exists a function $\eta \mapsto \lambda(\eta) < \infty$ such that if δ is sufficiently small as a function of η, then $\varrho(\mathscr{C}_k, \mathscr{C}_0) \le \lambda(\eta)$ for all $k \le N$. This is needed for $\eta = \eta(R)$ for all R in the compact set $[1, \bar{R}]$, so such a δ may be chosen as a function of $\bar{R}$ alone; conditions already imposed on δ above make it a function of both ε and $\bar{R}$.

In the region of all $x \in \mathbb{S}^2$ satisfying $\min(|x-z_0|, |x+z_0|) > Rr_0$, either $f_k \equiv 0$, or $\mathscr{C}_k$ has radius no less than $\frac{1}{4}Rr_0$, or the center z_k of $\mathscr{C}_k$ satisfies $\max(|z_k - z_0|, |z_k + z_0|) \ge \frac{1}{4}Rr_0$. Choose a function $R \mapsto \eta(R)$ that tends to 0 sufficiently slowly as $R \to \infty$ to ensure that $\lambda(\eta(R)) \to \infty$ sufficiently slowly that the latter two cases would contradict the inequality $\varrho(\mathscr{C}_k, \mathscr{C}_0) \le \lambda$, and therefore cannot arise. Then $\mathcal{F}(x) \equiv 0$ when $\min(|x-z_0|, |x+z_0|) > Rr_0$.

With the function η specified, Θ can be defined by

$$\Theta(R) = \gamma(\eta(R)). \tag{10-4}$$

Then (10-1) holds for all $R \in [1, \bar{R}]$.

We claim next that $\|\mathcal{F}\|_\infty < R|\mathscr{C}_0|^{-1/2}$ if R is sufficiently large as a function of η. Indeed, because the summands f_k have pairwise disjoint supports, it suffices to control $\max_{k \le N} \|f_k\|_\infty$. Again, by Lemma 8.4, $\|f_k\|_\infty \le C(\eta)|\mathscr{C}_k|^{-1/2}$. If $\eta(R)$ is chosen to tend to zero sufficiently slowly as $R \to \infty$ to ensure that $C(\eta(R))\lambda(\eta(R)) < R$ for all $k \le N$, then inequality (10-2) holds provided that Θ is defined by (10-4).

The final function η must be chosen to tend to zero slowly enough to satisfy the requirements of the proofs of both (10-1) and (10-2). □

11. Preliminaries for Step 7

Lemma 11.1. *Let $\Theta : [1, \infty) \to (0, \infty)$ satisfy $\Theta(R) \to 0$ as $R \to \infty$. Let $\delta > 0$. Then there exists $c > 0$ such that any nonnegative function $g \in L^2(\mathbb{R}^2)$ satisfying $\|g\|_2 = 1$ and the upper bounds*

$$\int_{|x| \ge R} g(x)^2\, dx + \int_{g(x) \ge R} g(x)^2\, dx \;\le\; \Theta(R) \quad \text{for all } R \ge 1,$$

has Fourier transform satisfying the lower bound

$$\int_{|\xi| \le \delta} |\hat{g}(\xi)|^2\, d\xi \ge c.$$

Proof. Let $g \in L^2(\mathbb{R}^2)$ satisfy the hypotheses. For $t > 0$, let $\varphi_t(y) = e^{-t|y|^2/2}$. Then

$$\int g\varphi_t\, dy = (2\pi)^{-2} \int \hat{g}(\xi)\widehat{\varphi_t}(\xi)\, d\xi = (2\pi)^{-1} t^{-1} \int \hat{g}(\xi) e^{-|\xi|^2/2t}\, d\xi.$$

For any R, $\rho \geq 1$, let $S = \{y : |y| \leq R \text{ and } g(y) \leq \rho\}$. Provided that R and ρ are chosen to be sufficiently large that $\Theta(R) + \Theta(\rho) \leq \frac{1}{2}$,

$$\int_{\mathbb{R}^2} g\varphi_t \, dy \geq e^{-tR^2/2} \int_S g(y)\, dy \geq e^{-tR^2/2} \rho^{-1} \int_S g^2(y)\, dy$$
$$= e^{-tR^2/2} \rho^{-1} \Big(\|g\|_2^2 - \int_{\mathbb{R}^2 \setminus S} g^2(y)\, dy \Big) \geq \tfrac{1}{2} e^{-tR^2/2} \rho^{-1}$$

for any $t > 0$. On the other hand, by the Cauchy–Schwarz inequality

$$\int_{|\xi| \geq \delta} |\hat{g}(\xi)|\, t^{-1} e^{-|\xi|^2/2t} \, d\xi \leq \pi^{1/2} t^{-1} \|\hat{g}\|_2 \Big(\int_{r=\delta}^{\infty} e^{-r^2/t} 2r \, dr \Big)^{1/2}$$
$$= \pi^{1/2} t^{-1} \Big(t \int_{s=\delta^2/t}^{\infty} e^{-s} \, ds \Big)^{1/2} = \pi^{1/2} t^{-1/2} e^{-\delta^2/2t}.$$

The Cauchy–Schwarz inequality also gives

$$\int_{|\xi| \leq \delta} |\hat{g}(\xi)|\, t^{-1} e^{-|\xi|^2/2t} \, d\xi \leq \Big(\int_{|\xi| \leq \delta} |\hat{g}(\xi)|^2 \, d\xi \Big)^{1/2} (2\pi)^{1/2} \Big(\int_0^{\infty} t^{-2} e^{-r^2/t}\, r \, dr \Big)^{1/2}$$
$$= \pi^{1/2} t^{-1/2} \Big(\int_{|\xi| \leq \delta} |\hat{g}(\xi)|^2 \, d\xi \Big)^{1/2}.$$

Therefore

$$\pi^{1/2} t^{-1/2} \Big(\int_{|\xi| \leq \delta} |\hat{g}(\xi)|^2 \, d\xi \Big)^{1/2} \geq \int_{\mathbb{R}^2} \hat{g}(\xi) t^{-1} e^{-|\xi|^2/2t} \, d\xi - \int_{|\xi| \geq \delta} |\hat{g}(\xi)|\, t^{-1} e^{-|\xi|^2/2t} \, d\xi$$
$$\geq \pi e^{-tR^2/2} \rho^{-1} - \pi^{1/2} t^{-1/2} e^{-\delta^2/2t}.$$

Now substitute $t = \delta^2/\gamma$, where $\gamma = \gamma(\delta) \geq 1$, to obtain

$$\pi^{1/2} \gamma^{1/2} \delta^{-1} \Big(\int_{|\xi| \leq \delta} |\hat{g}(\xi)|^2 \, d\xi \Big)^{1/2} \geq \pi e^{-\delta^2 R^2/2\gamma} \rho^{-1} - \pi^{1/2} \gamma^{1/2} \delta^{-1} e^{-\gamma/2}.$$

The quantities R and ρ have already been fixed, independent of δ. As δ also remains fixed while $\gamma \to \infty$, this last lower bound tends to $\pi\rho^{-1} - 0 > 0$. Thus choosing γ sufficiently large yields the desired lower bound. □

Lemma 11.2. *Let $c_0 > 0$. Let $\{g_\nu\}$ be any sequence of functions in $L^2(\mathbb{R}^2)$ satisfying $\|g_\nu\|_{L^2} = 1$ and $\int_{|\xi| \leq 1} |\widehat{g_\nu}(\xi)|^2 \, d\xi \geq c_0$. Then either there exists a function $\theta : [1, \infty) \to (0, \infty)$ satisfying*

$$\theta(s) \to 0 \quad \textit{as } s \to \infty$$

such that

$$\int_{|\xi| \geq s} |\widehat{g_\nu}(\xi)|^2 \, d\xi \leq \theta(s) \quad \textit{for all } s \in [1, \infty) \textit{ and all } \nu,$$

or there exist a subsequence $\nu_k \to \infty$ *and real constants* $\delta > 0$, $\varepsilon_k > 0$, *and* $S_k \ge s_k \ge 1$ *such that* $s_k \to \infty$, $\varepsilon_k \to 0$, $S_k = s_k^3$,

$$\int_{|\xi|\le s_k} |\widehat{g_{\nu_k}}(\xi)|^2\, d\xi \ge \delta, \quad \int_{|\xi|\ge S_k} |\widehat{g_{\nu_k}}(\xi)|^2\, d\xi \ge \delta, \quad \int_{s_k\le|\xi|\le S_k} |\widehat{g_{\nu_k}}(\xi)|^2\, d\xi < \varepsilon_k.$$

In this lemma, δ is permitted, in principle, to depend on $\{g_\nu\}$, and ε_k and s_k are permitted to depend on $\{g_\nu\}$ and on k in an arbitrary manner, provided only that they satisfy the stated conditions. The relation $S_k = s_k^3$ is chosen simply because it is convenient for the proof of Lemma 12.2 below; one could arrange to have S_k equal to any function of s_k that might be desired.

Proof. Define a sequence $\rho_1, \rho_2, \ldots$ by $\rho_1 = 2$ and by induction, $\rho_{j+1} = \rho_j^3$. If the first conclusion does not hold, then after passing to a subsequence and renumbering, we have

$$\int_{|\xi|\ge\rho_\nu} |\widehat{g_\nu}(\xi)|^2\, d\xi \ge \delta \quad \text{for all } \nu.$$

Consider a large ν. Since

$$\sum_{j=1}^{\nu-1} \int_{\rho_j\le|\xi|\le\rho_{j+1}} |\widehat{g_\nu}(\xi)|^2\, d\xi \le (2\pi)^2 \|g_\nu\|_2^2 \le (2\pi)^2$$

and there are $\nu - 1$ summands, there must exist $j(\nu)$ satisfying

$$\int_{\rho_j\le|\xi|\le\rho_{j+1}} |\widehat{g_\nu}(\xi)|^2\, d\xi \le C\nu^{-1}.$$

It suffices to set $s_\nu = \rho_{j(\nu)}$, $S_\nu = \rho_{j(\nu)+1} = s_\nu^3$, and $\varepsilon_\nu = C\nu^{-1}$. □

12. Step 7: Precompactness after rescaling

We begin the proof of Proposition 2.15. Let $\{f_\nu\}$ be as in Proposition 2.15. Set $g_\nu = \phi_\nu^*(f_\nu)$, where ϕ_ν is the rescaling map associated to $\mathscr{C}_\nu$. Let $r_\nu \to 0$. Then by definition of g_ν,

$$\|g_\nu\|_{L^2(\mathbb{R}^2)}^2 \to \tfrac{1}{2} \quad \text{as } \nu \to \infty,$$

so the results of the preceding section apply to $2^{1/2} g_\nu$, and hence to g_ν itself, uniformly in ν.

If the first alternative in the conclusion of Lemma 11.2 holds, then we obtain the conclusion of Proposition 2.15. Therefore we may assume, by passing to a subsequence, that $\{g_\nu\}$ satisfies the conclusions of the second alternative of Lemma 11.2.

Split

$$g_\nu = g_\nu^0 + g_\nu^\infty + g_\nu^\flat,$$

where

$$
\begin{aligned}
&\|g_\nu^0\|_2 \geq \delta, \qquad \widehat{g_\nu^0}(\xi) \text{ is supported where } |\xi| \leq 2s_\nu,\\
&\|g_\nu^\infty\|_2 \geq \delta, \qquad \widehat{g_\nu^\infty}(\xi) \text{ is supported where } |\xi| \geq \tfrac{1}{2}S_\nu,\\
&\|g_\nu^\flat\|_2 < \varepsilon_\nu,
\end{aligned}
$$

$$g_\nu^0, g_\nu^\infty \text{ are upper normalized with respect to } \mathcal{B}, \quad \text{and} \quad \varepsilon_\nu \to 0 \text{ as } \nu \to \infty.$$

Here $\delta > 0$ is a certain constant independent of ν, and $\mathcal{B}$ denotes the unit ball in $\mathbb{R}^2$. This splitting is accomplished via an appropriate C^∞ three term partition of unity in the Fourier space $\mathbb{R}^2_\xi$.

Write $\mathcal{C}_\nu = \mathcal{C}(z_\nu, r_\nu)$. The decomposition above of $g_\nu = \phi_\nu^*(f_\nu)$ induces a corresponding decomposition

$$f_\nu = F_\nu^0 + F_\nu^\infty + F_\nu^\flat,$$

where all three summands are real-valued and even and for all sufficiently large ν,

$$
\left.
\begin{aligned}
&F_\nu^0, F_\nu^\infty, F_\nu^\flat \text{ are upper even-normalized with respect to } \mathcal{C}_\nu,\\
&\|F_\nu^\flat\|_2 \to 0 \text{ as } \nu \to \infty, \quad \|F_\nu^0\|_2 \geq \delta/2, \quad \|F_\nu^\infty\|_2 \geq \delta/2,\\
&F_\nu^0 \text{ and } F_\nu^\infty \text{ are supported in } \mathcal{C}(z_\nu, \tfrac{1}{2}) \cup -\mathcal{C}(z_\nu, \tfrac{1}{2}).
\end{aligned}
\right\}
\tag{12-1}
$$

Moreover:

Lemma 12.1. *The decomposition $f_\nu = F_\nu^0 + F_\nu^\infty + F_\nu^\flat$ may be carried out so that the conditions above are satisfied, and moreover, for certain constants $C, C_N < \infty$, the summands F_ν^0 and F_ν^∞ are real-valued, even, and admit representations*

$$F_\nu^0(y) = \int_{H_\nu} a_\nu^{0,\pm}(\xi) e^{iy\cdot\xi}\, d\xi, \quad \textit{and} \quad F_\nu^\infty(y) = \int_{H_\nu} a_\nu^{\infty,\pm}(\xi) e^{iy\cdot\xi}\, d\xi, \tag{12-2}$$

where the representations with plus signs are valid for $y \in \mathcal{C}(z_\nu, \frac{1}{2})$, and those with minus signs are valid for $y \in -\mathcal{C}(z_\nu, \frac{1}{2})$, with Fourier coefficients $a_\nu^{0,\pm}$ and $a_\nu^{\infty,\pm}$ satisfying

$$\int_{r_\nu|\xi| \leq S_\nu/4} |a_\nu^{\infty,\pm}(\xi)|^2\, d\xi \leq C S_\nu^{-1} \qquad \textit{for all } \nu, \tag{12-3}$$

$$\int_{r_\nu|\xi| \geq 4s_\nu} |a_\nu^{0,\pm}(\xi)|^2\, d\xi \leq C_N s_\nu^{-N} \quad \textit{for all } \nu,\ \textit{for any } N < \infty. \tag{12-4}$$

Proof. By rotational symmetry, it suffices to prove this under the assumption that $z_\nu = (0, 0, 1)$ for all ν. Then $\phi_\nu^*(f_\nu)(x') = r_\nu f_\nu(r_\nu x', (1 - r_\nu^2|x'|^2)^{1/2})$ for $x' \in \mathbb{R}^2$, and $H_\nu = \{x = (x', 0) \in \mathbb{R}^2 \times \mathbb{R}^1\}$.

Once a representation of the required form is established for the restriction of f_ν to the hemisphere $\mathbb{S}^2_+ = \{y \in \mathbb{S}^2 : y_3 > 0\}$, the symmetry $f_\nu(-y) \equiv f_\nu(y)$ leads immediately to the desired representation for $y_3 < 0$. So we restrict attention to $\mathbb{S}^2_+$. For the remainder of this proof, we identify $(\xi', 0) \in \mathbb{R}^{2+1}$ with $\xi' \in \mathbb{R}^2$, and denote elements of $\mathbb{R}^2$ by ξ rather than by ξ'.

Fix a compactly supported C^∞ function $\zeta : \mathbb{R}^2 \to \mathbb{R}$ that is supported in $\{y' : |y'| < \frac{1}{2}\}$ and is $\equiv 1$ in $\{y' : |y'| \le \frac{1}{4}\}$. For $y' \in \mathbb{R}^2$, define

$$G_\nu^0(y') = (2\pi)^{-1}\zeta(y')r_\nu^{-1}\int_{\mathbb{R}^2} e^{ir_\nu^{-1}y'\cdot\xi}\,\widehat{g_\nu^0}(\xi)\,d\xi, \tag{12-5}$$

$$G_\nu^\infty(y') = (2\pi)^{-1}\zeta(y')r_\nu^{-1}\int_{\mathbb{R}^2} e^{ir_\nu^{-1}y'\cdot\xi}\,\widehat{g_\nu^\infty}(\xi)\,d\xi. \tag{12-6}$$

Then from the fact that g_ν^0 and g_ν^∞ are upper normalized with respect to $\mathcal{B}$, it follows that

$$\|r_\nu^{-1}g_\nu^0(r_\nu^{-1}\,\cdot) - G_\nu^0(\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \to 0 \quad \text{as } \nu \to \infty,$$

and likewise

$$\|r_\nu^{-1}g_\nu^\infty(r_\nu^{-1}\,\cdot) - G_\nu^\infty(\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \quad \text{as } \nu \to \infty,$$

using the hypothesis that $r_\nu \to 0$ coupled with the fact that the support of ζ is independent of r_ν. It can of course be arranged that G_ν^0 and G_ν^∞ are real-valued.

Define

$$F_\nu^\flat|_{\mathbb{S}_+^2}(y', y_3) = f_\nu(y', y_3) - G_\nu^0(y') - G_\nu^\infty(y').$$

where $y_3 = \sqrt{1-|y'|^2}$. The function $F_\nu^\flat|_{\mathbb{S}_+^2}$ is upper normalized with respect to $\mathscr{C}_\nu$ because all three summands in its definition are upper normalized. Since $\phi_\nu^*(f_\nu) = g_\nu^0 + g_\nu^\infty + g_\nu^\flat$, since $\|g_\nu^\flat\|_{L^2(\mathbb{R}^2)} \to 0$ as $\nu \to \infty$, since f_ν is upper normalized with respect to $\mathscr{C}_\nu$ and $r_\nu \to 0$, and since ϕ_ν^* is essentially an isometry from $L^2(\mathbb{S}_+^2)$ to $L^2(\mathbb{R}^2)$ for large ν (again because $r_\nu \to 0$), it follows that

$$\|F_\nu^\flat\|_{L^2(\mathbb{S}_+^2)} \to 0 \quad \text{as } \nu \to \infty.$$

When regarded in this way as functions of $y = (y', y_3) \in \mathbb{S}_+^2$, the summands $G_\nu^0(y')$ and $G_\nu^\infty(y')$ are each upper normalized with respect to the caps $\mathscr{C}_\nu$, because g_ν^0 and g_ν^∞ are upper normalized with respect to $\mathcal{B}$. It remains only to show that $G_\nu^0(y')$ can be represented in the form $\int_{\mathbb{R}^2} e^{iy'\cdot\xi}a_\nu^{0,+}(\xi)\,d\xi$, where $a_\nu^{0,+}$ satisfies the required bound (12-4), and likewise for G_ν^∞. To prove this for G_ν^0, it suffices to rewrite the product of $\zeta(y')$ with the inverse Fourier transform in (12-5) as the inverse Fourier transform of a convolution, and to combine the bound $|\widehat{\zeta}(\xi)| \le C_N(1+|\xi|)^{-N}$ for all N with the fact that $\widehat{g_\nu^0}(\xi) \equiv 0$ for $\{\xi : |\xi| > 2s_\nu\}$. The analysis of G_ν^∞ is essentially identical, using the given fact that $\widehat{g_\nu^\infty}(\xi) \equiv 0$ for $\{\xi : |\xi| < \frac{1}{2}S_\nu\}$. □

As $\nu \to \infty$,

$$\|f_\nu\sigma * f_\nu\sigma\|_2 \le \big\|(F_\nu^0\sigma * F_\nu^0\sigma) + (F_\nu^\infty\sigma * F_\nu^\infty\sigma)\big\|_2 + 2\|F_\nu^0\sigma * F_\nu^\infty\sigma\|_2 + o(1)$$

where $o(1)$ denotes a function that tends to zero as $\nu \to \infty$. Applying the triangle inequality to the first term does not lead to a useful bound. Instead,

$$\begin{aligned}\big\|(F_\nu^0\sigma * F_\nu^0\sigma) + (F_\nu^\infty\sigma * F_\nu^\infty\sigma)\big\|_2^2 \\ &\le \|F_\nu^0\sigma * F_\nu^0\sigma\|_2^2 + \|F_\nu^\infty\sigma * F_\nu^\infty\sigma\|_2^2 + 2|\langle F_\nu^0\sigma * F_\nu^0\sigma,\ F_\nu^\infty\sigma * F_\nu^\infty\sigma\rangle| \\ &= \|F_\nu^0\sigma * F_\nu^0\sigma\|_2^2 + \|F_\nu^\infty\sigma * F_\nu^\infty\sigma\|_2^2 + 2|\langle F_\nu^0\sigma * F_\nu^\infty\sigma,\ F_\nu^0\sigma * F_\nu^\infty\sigma\rangle|\end{aligned}$$

since F_ν^0 and F_ν^∞ are real and even. Therefore*and* since F_ν^0, F_ν^∞ have uniformly bounded L^2 norms,

$$\|f_\nu\sigma * f_\nu\sigma\|_2^2 \le \|F_\nu^0\sigma * F_\nu^0\sigma\|_2^2 + \|F_\nu^\infty\sigma * F_\nu^\infty\sigma\|_2^2 + C\|F_\nu^0\sigma * F_\nu^\infty\sigma\|_2^2 + o(1). \tag{12-7}$$

The following key lemma will be proved below, in Section 14.

Lemma 12.2. *Let F_ν^0 and F_ν^∞ be upper even-normalized with respect to a sequence of caps of radii $r_\nu \to 0$. Assume that F_ν^0 and F_ν^∞ admit Fourier representations satisfying the inequalities specified in Lemma 12.1. Then*

$$\|F_\nu^0\sigma * F_\nu^\infty\sigma\|_{L^2(\mathbb{R}^3)} \to 0.$$

Corollary 12.3. *The second alternative in the conclusion of Lemma 11.2 cannot hold.*

Proof. Assume Lemma 12.2. Then by (12-7),

$$\|f_\nu\sigma * f_\nu\sigma\|_2^2 \le \|F_\nu^0\sigma * F_\nu^0\sigma\|_2^2 + \|F_\nu^\infty\sigma * F_\nu^\infty\sigma\|_2^2 + o(1) \le \mathbf{S}^4\|F_\nu^0\|_2^4 + \mathbf{S}^4\|F_\nu^\infty\|_2^4 + o(1).$$

Since $S_\nu/s_\nu \to \infty$ and $\|F_\nu^\flat\|_2 \to 0$, it follows easily from (12-3) and (12-4) that

$$\|F_\nu^0\|_2^2 + \|F_\nu^\infty\|_2^2 \le (1+o(1))\|f_\nu\|_2^2 = 1 + o(1).$$

Since $\min(\|F_\nu^0\|_2, \|F_\nu^\infty\|_2) \ge \delta/2$, this forces

$$\max(\|F_\nu^0\|_2^2, \|F_\nu^\infty\|_2^2) \le 1 - \rho$$

for all sufficiently large ν, for some $\rho > 0$ independent of ν. It follows that

$$\begin{aligned}\mathbf{S}^4\|F_\nu^0\|_{L^2(\sigma)}^4 + \mathbf{S}^4\|F_\nu^\infty\|_{L^2(\sigma)}^4 &\le \mathbf{S}^4\big(\|F_\nu^0\|_{L^2(\sigma)}^2 + \|F_\nu^\infty\|_{L^2(\sigma)}^2\big)\max(\|F_\nu^0\|_2^2, \|F_\nu^\infty\|_2^2) \\ &\le \mathbf{S}^4(1+o(1))(1-\rho).\end{aligned}$$

We conclude that

$$\limsup_{\nu\to\infty} \|f_\nu\sigma * f_\nu\sigma\|_{L^2(\mathbb{R}^3)}^2 < \mathbf{S}^4,$$

contradicting the assumption that $\{f_\nu\}$ was an extremizing sequence. □

Combining the results above, the proof of Proposition 2.15 is complete except for the proof of Lemma 12.2.

13. Step 8: Excluding small caps

In this section we prove Proposition 2.16, which asserts that the radii r_ν of the caps $\mathscr{C}_\nu$ associated to an extremizing sequence $\{f_\nu\}$ of positive even functions cannot tend to zero.

Lemma 13.1. *Let $\{f_\nu\}$ be any sequence of real-valued, even functions on $\mathbb{S}^2$ satisfying $\|f_\nu\|_{L^2} = 1$. Suppose that f_ν is upper even-normalized with respect to a cap $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$, uniformly in ν. Suppose that the sequence of pullbacks $\phi_\nu^*(f_\nu)$ satisfies the first alternative in the conclusion of Lemma 11.2. Suppose that $r_\nu \to 0$. Then there exists a sequence of functions $F_\nu : \mathbb{P}^2 \to \mathbb{R}$ satisfying $\|F_\nu\|_2 \to 1$ such that*

$$\limsup_{\nu\to\infty} \|F_\nu\sigma_P * F_\nu\sigma_P\|_2 \geq (3/2)^{-1/2} \limsup_{\nu\to\infty} \|f_\nu\sigma * f_\nu\sigma\|_2.$$

Proof of Proposition 2.16. Let $\{f_\nu\}$ be an extremizing sequence of nonnegative even functions for the inequality (2-1) satisfying $\|f_\nu\|_2 = 1$. There exists a sequence of caps $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$ such that each f_ν is upper even-normalized with respect to $\mathscr{C}_\nu$. We must prove that $\inf_\nu r_\nu > 0$.

If not, then by passing to a subsequence we may assume that $r_\nu \to 0$. By Proposition 2.15, the sequence of pullbacks $g_\nu = \phi_\nu^*(f_\nu)$ is precompact in $L^2(\mathbb{R}^2)$. Thus the hypotheses of Lemma 13.1 are satisfied, so there exists a sequence of functions $F_\nu \in L^2(\mathbb{P}^2)$ satisfying its conclusions.

Now $\|F_\nu\sigma_P * F_\nu\sigma_P\|_2 \leq \mathbf{P}^2 \|F_\nu\|^2_{L^2(\mathbb{P}^2)}$ by the definition of $\mathbf{P}$. Consequently

$$\limsup_{\nu\to\infty} \|f_\nu\sigma * f_\nu\sigma\|_2 \leq (3/2)^{1/2}\mathbf{P}^2.$$

The left side tends to $\mathbf{S}^2$ since $\{f_\nu\}$ is an extremizing sequence for (2-1), so $\mathbf{S}^2 \leq (3/2)^{1/2}\mathbf{P}^2$, contradicting the inequality $\mathbf{S} \geq 2^{1/4}\mathbf{P}$ of Lemma 2.4. □

Proof of Lemma 13.1. Write $\mathscr{C}_\nu = \mathscr{C}(z_\nu, r_\nu)$. Decompose $2^{1/2} f_\nu(x) = f_\nu^+(x) + f_\nu^+(-x) + f_\nu^\flat(x)$, where f_ν^+ is real, f_ν^+ is supported in $\mathscr{C}(z_\nu, r_\nu^{1/2})$, $\|f_\nu^\flat\|_2 \to 0$, and the functions $\phi_\nu^*(f_\nu^+)$ satisfy the first alternative of the conclusions of Lemma 11.2, uniformly in ν.

Since f_ν is even and $\|f_\nu\|_2 = 1$, we have $\|f_\nu^+\|_2 \to 1$ as $\nu \to \infty$. Moreover $g_\nu(x) = f_\nu^+(x) + f_\nu^+(-x)$ satisfies

$$\|g_\nu\sigma * g_\nu\sigma\|_2^2 / \|g_\nu\|_2^4 \equiv \tfrac{3}{2} \big\| f_\nu^+\sigma * f_\nu^+\sigma \big\|_2^2 \Big/ \big\| f_\nu^+ \big\|_2^4,$$

and therefore

$$\limsup_{\nu\to\infty} \|f_\nu^+\sigma * f_\nu^+\sigma\|_2^2 = (3/2)^{-1} \limsup_{\nu\to\infty} \|f_\nu\sigma * f_\nu\sigma\|_2^2.$$

By rotation symmetry, we may suppose that $z_\nu = (0, 0, 1)$ for all ν. Define $F_\nu : \mathbb{P}^2 \to [0, \infty)$ by

$$F_\nu(y, |y|^2/2) = r_\nu f_\nu^+\big(r_\nu y, (1 - r_\nu^2|y|^2)^{1/2}\big)$$

for $y \in \mathbb{R}^2$. The function F_ν will also be regarded as an element of $L^2(\mathbb{R}^2, dy)$ by $F_\nu(y) = F_\nu(y, |y|^2/2)$. Then $\|F_\nu\|_{L^2(\mathbb{P}^2, \sigma_P)} = \|F_\nu\|_{L^2(\mathbb{R}^2)} \to 1$ as $\nu \to \infty$.

It remains to prove that

$$\limsup_{\nu\to\infty} \|\widehat{F_\nu\sigma_P}\|^4_{L^4(\mathbb{R}^3)} \geq \limsup_{\nu\to\infty} \|\widehat{f_\nu^+\sigma}\|^4_{L^4(\mathbb{R}^3)}.$$

We have

$$\int_{|y|\ge R} F_\nu(y)^2\,dy+\int_{F_\nu(y)\ge R} F_\nu(y)^2\,dy+\int_{|\xi|\ge R}|\widehat{F_\nu}(\xi)|^2\,d\xi\to 0 \tag{13-1}$$

as $R\to\infty$, uniformly in ν.

Thus we must compare $\widehat{F_\nu\sigma_P}(x,t)=\int e^{-ix\cdot y-it|y|^2/2}F_\nu(y)\,dy$ with

$$\begin{aligned}\widehat{f_\nu^+\sigma}(x,t)&=\int_{\mathbb{R}^2} e^{-ix\cdot v-it(1-|v|^2)^{1/2}} f_\nu^+(v,(1-|v|^2)^{1/2})\,d\sigma(v,(1-|v|^2)^{1/2})\\&=\int_{\mathbb{R}^2} e^{-ix\cdot v-it(1-|v|^2)^{1/2}} f_\nu^+(v,(1-|v|^2)^{1/2})\,(1-|v|^2)^{-1/2}\,dv.\end{aligned}$$

In the latter integral, substitute $v=r_\nu y$ to obtain

$$\begin{aligned}&r_\nu^{-1}\widehat{f_\nu^+\sigma}(r_\nu^{-1}x,-r_\nu^{-2}t)\\&\quad=r_\nu^{-1}r_\nu^2\int_{\mathbb{R}^2} e^{-ix\cdot y+itr_\nu^{-2}(1-r_\nu^2|y|^2)^{1/2}} f_\nu^+(r_\nu y,(1-r_\nu^2|y|^2)^{1/2})\,(1-r_\nu^2|y|^2)^{-1/2}\,dy\\&\quad=\int_{\mathbb{R}^2} e^{-ix\cdot y+itr_\nu^{-2}(1-r_\nu^2|y|^2)^{1/2}} F_\nu(y)(1-r_\nu^2|y|^2)^{-1/2}\,dy\\&\quad=e^{itr_\nu^{-2}}\int_{\mathbb{R}^2} e^{-ix\cdot y-it|y|^2/2}F_\nu(y)h_\nu(t,y)\,dy,\end{aligned}$$

where

$$h_\nu(t,y)=e^{it\psi_\nu(y)}(1-r_\nu^2|y|^2)^{-1/2}\quad\text{and}\quad\psi_\nu(y)=-r_\nu^{-2}+|y|^2/2+r_\nu^{-2}(1-r_\nu^2|y|^2)^{1/2}.$$

Thus

$$\|\widehat{f_\nu^+\sigma}\|_4^4=\int_{\mathbb{R}}\int_{\mathbb{R}^2}\left|r_\nu^{-1}\widehat{f_\nu^+\sigma}(r_\nu^{-1}x,-r_\nu^{-2}t)\right|^4dx\,dt=\left\|\int_{\mathbb{R}^2} e^{-ix\cdot y-it|y|^2/2}F_\nu(y)h_\nu(t,y)\,dy\right\|_{L^4(\mathbb{R}^3)}^4.$$

It will be important that on any compact subset of $\mathbb{R}^1_t\times\mathbb{R}^2_y$,

$$h_\nu(t,y)\to 1 \text{ in the } C^N \text{ norm as } \nu\to\infty, \text{ for all } N<\infty. \tag{13-2}$$

Define

$$u_\nu(x,t)=\int_{\mathbb{R}^2} e^{-ix\cdot y-it|y|^2/2}F_\nu(y)h_\nu(t,y)\,dy\quad\text{and}\quad\tilde u_\nu(x,t)=\int e^{-ix\cdot y-it|y|^2/2}F_\nu(y)\,dy.$$

Lemma 13.2. *We have*

$$\int_{|(x,t)|\ge R|} u_\nu(x,t)|^4\,dx\,dt\to 0\quad \textit{as } R\to\infty, \textit{uniformly in } \nu,$$

$$\int_{|(x,t)|\ge R} |\tilde u_\nu(x,t)|^4\,dx\,dt\to 0\quad \textit{as } R\to\infty, \textit{uniformly in } \nu.$$

Proof. Define operators T_ν and T from $L^2(\mathbb{R}^2)$ to $L^4(\mathbb{R}^3)$ by

$$T_\nu g(x,t)=\int_{\mathbb{R}^2} e^{-ix\cdot y-it|y|^2/2}g(y)\chi_{r_\nu^{-1}|y|\le 1/2}(y)h_\nu(t,y)\,dy,\ Tg(x,t)=\int e^{-ix\cdot y-it|y|^2/2}g(y)\,dy.$$

The operator $T : L^2(\mathbb{R}^2) \to L^4(\mathbb{R}^3)$ is bounded. Although the operators T_ν are written in coordinates that disguise this fact, they are bounded from $L^2(\mathbb{R}^2)$ to $L^4(\mathbb{R}^3)$ uniformly in ν, being obtained via norm-preserving changes of variables from the single bounded operator $L^2(\mathbb{S}^2, \sigma) \ni h \mapsto \widehat{h\sigma}$.

If $g \in C^2(\mathbb{R}^2)$ has compact support, then $|T_\nu g(x,t)| \le C_g |(x,t)|^{-1}$, where C_g depends only on the C^1 norm of g and on the diameter of its support, provided that ν is sufficiently large that the support of g is contained in $B(0, r_\nu^{-1})$. This follows from (13-2) together with the method of stationary phase; the phase functions appearing in the definition of T_ν have uniformly nondegenerate critical points (if any), uniformly in ν.

These two facts, together with the three uniform inequalities (13-1), lead directly to the stated conclusion for u_ν by a routine argument.

A slightly simpler application of the same reasoning applies to $\tilde{u}_\nu$. □

Therefore it suffices to prove that for any $R < \infty$,

$$\int_{|(x,t)| \le R} \left| u_\nu(x,t) - \tilde{u}_\nu(x,t) \right|^4 dx\, dt \to 0 \quad \text{as } \nu \to \infty. \tag{13-3}$$

If $g \in L^1$ has compact support, then

$$|T_\nu(g)(x,t) - T(g)(x,t)| \to 0, \text{ uniformly for all } |(x,t)| \le R. \tag{13-4}$$

Since T_ν and T are uniformly bounded operators from L^2 to L^4, and since the class of all compactly supported $g \in L^1$ is dense in L^2, (13-3) follows from (13-4). □

14. Estimation of the cross term $\| F_\nu^0 \sigma * F_\nu^\infty \sigma \|_2^2$

To prove Lemma 12.2, let f_ν, F_ν^0 and F_ν^∞ be as above. Let f_ν be upper even-normalized with respect to a cap $\mathscr{C}_\nu$ of radius r_ν. Since the inequality in question is invariant under rotations of $\mathbb{R}^3$, we may suppose without loss of generality that $\mathscr{C}_\nu$ is centered at the north pole $z_0 = (0,0,1)$.

Decompose $F_\nu^0 = F_\nu^{0,+} + F_\nu^{0,-}$, where both summands are real-valued, $F_\nu^{0,+}$ is supported in $\mathscr{C}(z_0, \frac{1}{2})$, $F_\nu^{0,-}(x) = F_\nu^{0,+}(-x)$, $F_\nu^{0,\pm}$ is upper normalized with respect to $\mathscr{C}(\pm z_0, r_\nu)$, and $F_\nu^{0,\pm}$ have the same Fourier representations (12-2) as F_ν^0. There is a parallel decomposition $F_\nu^\infty = F_\nu^{\infty,+} + F_\nu^{\infty,-}$. By Lemma 3.2,

$$\| F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma \|_2 = \| F_\nu^{0,-}\sigma * F_\nu^{\infty,-}\sigma \|_2 = \| F_\nu^{0,-}\sigma * F_\nu^{\infty,+}\sigma \|_2 = \| F_\nu^{0,+}\sigma * F_\nu^{\infty,-}\sigma \|_2.$$

Therefore it suffices to bound $\| F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma \|_2$.

Lemma 14.1. *Let $\delta_\nu, \delta_\nu^* > 0$ be sequences of positive numbers that satisfy*

$$\delta_\nu / r_\nu^2 \to 0 \quad \text{and} \quad \delta_\nu^* / r_\nu^2 \to \infty.$$

Then, with $A := \{x \in \mathbb{R}^3 : |x| > 2 - \delta_\nu \text{ or } |x| < 2 - \delta_\nu^*\}$,

$$\| F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma \|_{L^2(A)} \to 0 \quad \text{as } \nu \to \infty.$$

Proof. Since $F_\nu^{0,+}$ and $F_\nu^{\infty,+}$ are upper normalized with respect to $\mathscr{C}_\nu$, Corollary 7.3 asserts that the region $|x| > 2-\delta_\nu$ makes a small contribution for large ν. To handle the region $|x| < 2-\delta_\nu^*$, choose a sequence $t_\nu \geq 1$ tending slowly to infinity. Decompose $F_\nu^{0,+} = F_\nu^{0,+}\chi_{\mathscr{C}(z_0,t_\nu r_\nu)} + F_\nu^{0,+}\chi_{\mathbb{S}^2\setminus\mathscr{C}(z_0,t_\nu r_\nu)}$, and decompose $F_\nu^{\infty,+}$ in the same way. If $t_\nu \to \infty$ sufficiently slowly, then the main term $F_\nu^{0,+}\chi_{\mathscr{C}(z_0,t_\nu r_\nu)}\sigma * F_\nu^{\infty,+}\chi_{\mathscr{C}(z_0,t_\nu r_\nu)}\sigma$ is supported where $|x| > 2-\delta_\nu^*$. Expanding $F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma$ according to this decomposition leaves three more terms. Each of these has small norm in $L^2(\mathbb{R}^3)$ for large ν, because $\|F_\nu^{0,+}\|_{L^2(\mathbb{S}^2\setminus\mathscr{C}(z_0,t_\nu r_\nu))} \to 0$ and $\|F_\nu^{\infty,+}\|_{L^2(\mathbb{S}^2\setminus\mathscr{C}(z_0,t_\nu r_\nu))} \to 0$. □

If h_1 and h_2 are supported in $\mathscr{C}(z_0, r)$, then $h_1\sigma * h_2\sigma$ is supported in $\{x \in \mathbb{R}^3 : |x-2z_0| \leq Cr\}$. Since $F_\nu^{0,+}$ and $F_\nu^{\infty,+}$ are upper normalized with respect to $\mathscr{C}(z_\nu, r_\nu)$, and since $r_\nu \to 0$, it follows from the inequality $\|h_1\sigma * h_2\sigma\|_{L^2(\mathbb{R}^3)} \leq C\|h_1\|_2\|h_2\|_2$ that

$$\int_{|x-2z_0|\geq 1/100} |(F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma)(x)|^2\,dx \to 0 \quad \text{as } \nu\to\infty.$$

On the other hand, if $|x-2z_0| \leq 1/100$, then for all sufficiently large ν, $(F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma)(x)$ depends only on the restrictions of $F_\nu^{0,+}$ and $F_\nu^{\infty,+}$ to $\mathscr{C}(z_0, 1/10)$. This has the following significance in terms of the Fourier representations (12-3), (12-4) of Lemma 12.1:

$$F_\nu^{0,+}(x) = \int_{r_\nu|\zeta|\leq 4s_\nu} e^{ix\zeta} a_\nu^{0,+}(\zeta)\,d\zeta + o(1) \quad \text{in } L^2(\mathscr{C}(z_0, 1/10)) \text{ as } \nu\to\infty \tag{14-1}$$

by virtue of (12-4); this does not follow for $L^2(\mathbb{S}^2)$ because surface measure on $\mathbb{S}^2$ is not approximately equivalent to Lebesgue measure on $\{(x_1, x_2, 0)\}$ near the equator $\{x \in \mathbb{S}^2 : x_3 = 0\}$. Likewise, by (12-3),

$$F_\nu^{\infty,+}(x) = \int_{r_\nu|\zeta|\geq S_\nu/4} e^{ix\zeta} a_\nu^{\infty,+}(\zeta)\,d\zeta + o(1) \quad \text{in } L^2(\mathscr{C}(z_0, 1/10)) \text{ as } \nu\to\infty. \tag{14-2}$$

Henceforth we simplify notation by writing a_ν^0 in place of $a_\nu^{0,+}$ and a_ν^∞ in place of $a_\nu^{\infty,+}$, and we will take these functions to be supported in the sets $r_\nu|\zeta| \leq 4s_\nu$ and $r_\nu|\zeta| \geq S_\nu/4$, respectively.

Set $H = \{\xi \in \mathbb{R}^3 : \xi_3 = 0\}$, and identify $(\xi_1, \xi_2, 0) \in H$ with $(\xi_1, \xi_2) \in \mathbb{R}^2$. Denote a region $\mathscr{A}_\nu$ and an interval I_ν by

$$\mathscr{A}_\nu = \{x \in \mathbb{R}^3 : 2-\delta_\nu^* \leq |x| \leq 2-\delta_\nu \text{ and } |x-2z_0| < 1/100\} \quad \text{and} \quad I_\nu = [2-\delta_\nu^*, 2-\delta_\nu].$$

It remains only to estimate $\|F_\nu^{\infty,+}\sigma * F_\nu^{0,+}\sigma\|_{L^2(\mathscr{A}_\nu)}$. For $x \in \mathscr{A}_\nu$ and for all sufficiently large ν, $(F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma)(x)$ depends only on the restrictions of $F_\nu^{0,+}$, $F_\nu^{\infty,+}$ to $\mathscr{C}(z_0, 1/10)$. Therefore in majorizing $\|F_\nu^{\infty,+}\sigma * F_\nu^{0,+}\sigma\|_{L^2(\mathscr{A}_\nu)}$, we may replace $F_\nu^{0,+}(x)$ by $\int_{r_\nu|\zeta|\leq 4s_\nu} e^{ix\zeta}a_\nu^0(\zeta)\,d\zeta$ and $F_\nu^{\infty,+}(x)$ by $\int_{r_\nu|\zeta|\geq S_\nu/4} e^{ix\zeta}a_\nu^\infty(\zeta)\,d\zeta$, at the expense of additional terms that are $o(1)$ as $\nu\to\infty$. We will continue to denote these modified functions by $F_\nu^{0,+}$, $F_\nu^{\infty,+}$.

Set $h_\zeta = e_{-i\zeta}F_\nu^{\infty,+}$ for $r_\nu|\zeta| \leq 4s_\nu$. Let

$$H^* = \{\zeta \in H : r_\nu|\zeta| \leq 4s_\nu\}.$$

By (7-4), (7-5), (14-1), and (14-2),

$$\begin{aligned}\|F_\nu^{\infty,+}\sigma * F_\nu^{0,+}\sigma\|_{L^2(\mathcal{A}_\nu)}^2 &\le C\int_{I_\nu}\Big\|\int_{H^*}|a_\nu^0(\zeta)|\cdot|T_{\rho(t)}h_\zeta|\,d\zeta\Big\|_{L^2(\mathbb{S}^2)}^2 dt + o(1)\\ &\le C\int_{I_\nu}\Big(\int_{H^*}|a_\nu^0(\zeta)|\cdot\|T_{\rho(t)}h_\zeta\|_{L^2(\mathbb{S}^2)}\,d\zeta\Big)^2 dt + o(1)\\ &\le C\|a_\nu^0\|_2^2\int_{H^*}\int_{I_\nu}\|T_{\rho(t)}h_\zeta\|_{L^2(\mathbb{S}^2)}^2\,dt\,d\zeta + o(1)\\ &\le C\int_{H^*}\int_{I_\nu}\|T_{\rho(t)}h_\zeta\|_{L^2(\mathbb{S}^2)}^2\,dt\,d\zeta + o(1)\end{aligned}$$

by the Minkowski and Cauchy–Schwarz inequalities. Inserting the Fourier integral operator bound $\|T_\rho(h_\zeta)\|_2^2 \le C\|(I-\rho^2\Delta)^{-1/4}h_\zeta\|_2^2$ yields

$$\begin{aligned}\|F_\nu^{\infty,+}\sigma * F_\nu^{0,+}\sigma\|_{L^2(\mathcal{A}_\nu)}^2 &\le C\int_{\zeta\in H^*}\int_{I_\nu}\int_{\xi\in H}(1+\rho(t)|\xi|)^{-1}|\widehat{h_\zeta(\xi)}|^2\,d\xi\,dt\,d\zeta + o(1)\\ &= C\int_{\zeta\in H^*}\int_{I_\nu}\int_{\xi\in H}(1+\rho(t)|\xi|)^{-1}|a_\nu^\infty(\xi-\zeta)|^2\,d\xi\,dt\,d\zeta\\ &\sim s_\nu^2 r_\nu^{-2}\int_{I_\nu}\int_H(1+\rho(t)|\xi|)^{-1}|a_\nu^\infty(\xi)|^2\,d\xi\,dt + o(1),\end{aligned} \tag{14-3}$$

since $|\xi|\gg|\zeta|$ for ζ in the support of a_ν^0 and ξ in the support of a_ν^∞. Next,

$$\begin{aligned}\int_H(1+\rho|\xi|)^{-1}|a_\nu^\infty(\xi)|^2\,d\xi &\le C\int_{r_\nu|\xi|\le c_0S_\nu}|a_\nu^\infty(\xi)|^2\,d\xi + C\int_{r_\nu|\xi|\ge c_0S_\nu}(1+\rho|\xi|)^{-1}|a_\nu^\infty(\xi)|^2\,d\xi\\ &\le CS_\nu^{-1}\|F_\nu^{\infty,+}\|_2^2 + C\max_{r_\nu|\xi|\ge c_0S_\nu}(1+\rho|\xi|)^{-1}\cdot\|F_\nu^{\infty,+}\|_2^2\\ &\le CS_\nu^{-1} + C\rho^{-1}r_\nu S_\nu^{-1}.\end{aligned}$$

The first term after the first inequality was estimated using (12-3). Inserting the final line into (14-3) yields

$$\begin{aligned}\|F_\nu^{\infty,+}\sigma * F_\nu^{0,+}\sigma\|_{L^2(\mathcal{A}_\nu)}^2 &\le Cs_\nu^2r_\nu^{-2}\int_{I_\nu}\big(S_\nu^{-1}+\rho(t)^{-1}r_\nu S_\nu^{-1}\big)\,dt\\ &\le Cs_\nu^2r_\nu^{-2}\int_{I_\nu}(S_\nu^{-1}+(2-t)^{-1/2}r_\nu S_\nu^{-1})\,dt\\ &\qquad\text{since } (t/2)^2+\rho(t)^2=1 \text{ implies } \rho(t)\ge C(2-t)^{1/2}\\ &= Cs_\nu^2S_\nu^{-1}r_\nu^{-2}\int_{I_\nu}(1+r_\nu(2-t)^{-1/2})\,dt\\ &\le Cs_\nu^2S_\nu^{-1}r_\nu^{-2}|I_\nu|\big(1+\max_{t\in I_\nu}r_\nu(2-t)^{-1/2}\big)\\ &\le Cs_\nu^2S_\nu^{-1}(r_\nu^{-2}\delta_\nu^*)(1+\delta_\nu^{-1/2}r_\nu)\le Cs_\nu^{-1}(r_\nu^{-2}\delta_\nu^*)(1+\delta_\nu^{-1/2}r_\nu)\end{aligned}$$

since $S_\nu\ge s_\nu^3$.

Combining all terms, we have shown that

$$\| F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma \|_2^2 \le o(1) + C s_\nu^{-1}(r_\nu^{-2}\delta_\nu^*)(1+\delta_\nu^{-1/2} r_\nu)$$

as $\nu \to \infty$, provided that $\delta_\nu / r_\nu^2 \to 0$ and $\delta_\nu^*/r_\nu^2 \to \infty$. Since $s_\nu \to \infty$, it is possible to choose δ_ν and δ_ν^* to satisfy the additional constraint

$$s_\nu^{-1}(r_\nu^{-2}\delta_\nu^*)\big(1+\delta_\nu^{-1/2} r_\nu\big) \to 0 \quad \text{as } \nu \to \infty.$$

With such a choice, we obtain

$$\| F_\nu^{0,+}\sigma * F_\nu^{\infty,+}\sigma \|_2^2 \to 0 \quad \text{as } \nu \to \infty,$$

completing the proof of Lemma 12.2. □

15. Step 9: Large caps

We now prove Proposition 2.17. The proof is quite similar to that of Proposition 2.15, without the complication of ensuring uniformity of bounds as $r_\nu \to 0$. However, the proof of Proposition 2.15 also exploited the condition $r_\nu \to 0$ in a positive way, and substantive modification is therefore required here. Matters here that are essentially identical to corresponding matters in the earlier proof will be treated sketchily.

There is given an extremizing sequence $\{f_\nu\}$ of even nonnegative functions satisfying $\|f_\nu\|_{L^2(\mathbb{S}^2)} = 1$, each of which is upper even-normalized with respect to a certain cap $\mathscr{C}(z_\nu, r_\nu)$. It is given that $r_* = \inf_\nu r_\nu$ is strictly positive.

Introduce a C^∞ partition of unity of $\mathbb{S}^2$ by nonnegative functions η_j, each of which is supported in a cap $\mathscr{C}(z_j, \frac{1}{2})$. The points z_j and functions η_j are to be chosen independent of ν. Let $\phi_j : \mathbb{R}^2 \to \mathbb{S}^2$ and $\phi_j^* : L^2(\mathbb{S}^2) \to L^2(\mathbb{R}^2)$ be the associated mappings.

Since $r_* \le r_\nu \le 1$, the uniform upper normalization of f_ν means simply that $\|f_\nu\|_{L^2(\mathbb{S}^2)} \le 1$, and there exists a function Θ that is independent of ν and satisfies $\Theta(R) \to \infty$ as $R \to \infty$, such that

$$\int_{|f_\nu(x)|\ge R} |f_\nu(x)|^2 \, d\sigma(x) \le \Theta(R) \quad \text{for all } \nu.$$

Thus the radii r_ν no longer enter into the discussion.

Decompose $f_\nu = \sum_j f_{\nu,j}$, where $f_{\nu,j} = \eta_j f_\nu$. By identifying the plane tangent to $\mathbb{S}^2$ at z_j with a fixed copy of $\mathbb{R}^2$, we may regard each $g_{\nu,j} = \phi_j^*(f_{\nu,j})$ as an element of $L^2(\mathbb{R}^2)$; thus the functions $g_{\nu,j}$, and hence their Fourier transforms, have a common domain. The functions $g_{\nu,j}$ are supported in $\{y \in \mathbb{R}^2 : |y| \le \frac{1}{2}\}$, and again $\int_{|g_{\nu,j}(y)|\ge R} |g_{\nu,j}(y)|^2 \, dy \le \Theta(R)$, where $\Theta(R) \to 0$ as $R \to \infty$.

The analogue of Lemma 11.2 in this simplified situation is the following dichotomy: Either there exists a function $\theta : [1,\infty) \to (0,\infty)$ satisfying $\theta(s) \to 0$ as $s \to \infty$ such that

$$\int_{|\xi|\ge s} \sum_j |\widehat{g_{\nu,j}}(\xi)|^2 \, d\xi \le \theta(s) \quad \text{for all } s \in [1,\infty) \text{ and all } \nu, \tag{15-1}$$

or there exist $\delta, \varepsilon_\nu, s_\nu, S_\nu$ as in Lemma 11.2 such that the conclusions in the second case of that lemma hold, with $|\widehat{g_\nu}|^2$ replaced by $\sum_j |\widehat{g_{\nu,j}}|^2$. The proof of this dichotomy is essentially identical to the proof of Lemma 11.2 itself.

If (15-1) holds, then the conclusion of Proposition 2.17 is simply a reformulation of the conjunction of the upper normalization bounds for f_ν with (15-1); the desired decomposition of f_ν is obtained by expressing each $g_{\nu,j}$ as an inverse Fourier transform, splitting the resulting integral with respect to ξ into large $|\xi|$ and smaller $|\xi|$ regions, and reversing the mapping ϕ_j^* to transplant both summands to $\mathbb{S}^2$. The contribution of sufficiently large $|\xi|$ will have small $L^2(\mathbb{S}^2)$ norm, while the contribution of smaller $|\xi|$ will satisfy an adequate C^1 norm bound.

It remains only to demonstrate that the second case of the dichotomy cannot arise; there cannot exist $\delta, \varepsilon_\nu, s_\nu, S_\nu$ satisfying all conclusions of the second case of Lemma 11.2. Suppose to the contrary that this situation were to arise. Denote by $\phi_{j,*}^{-1}$ the left inverse of ϕ_j^*, mapping functions supported in $\{y \in \mathbb{R}^2 : |y| \le \frac{3}{4}\}$ to functions supported in $\mathscr{C}(z_j, \frac{3}{4}) \subset \mathbb{S}^2$. By summing over j, one would obtain as in (12-1) a decomposition

$$f_\nu = F_\nu^0 + F_\nu^\infty + F_\nu^\flat, \tag{15-2}$$

where $\lim_{\nu\to\infty} \|F_\nu^\flat\|_2 = 0$, F_ν^∞ is highly oscillatory, and F_ν^0 is slowly varying in comparison with F_ν^∞. Here for instance

$$F_\nu^\infty = \sum_j \phi_{j,*}^{-1}(\zeta \cdot g_{\nu,j}^\infty), \quad \text{where } \widehat{g_{\nu,j}^\infty}(\xi) = (1 - m(\xi/S_\nu))\widehat{g_{\nu,j}}(\xi)$$

and where the C^∞ cutoff functions ζ and m have the following properties: $\zeta \in C^\infty(\mathbb{R}^2)$ is $\equiv 1$ on the ball $B(0, \frac{5}{8})$, and is supported on $B(0, \frac{3}{4})$, while $m(\xi) \equiv 0$ for $|\xi| \ge \frac{3}{8}$ and $m(\xi) \equiv 1$ for $|\xi| \le \frac{1}{4}$. F_ν^0 is defined in the same way, with $1 - m(\xi/S_\nu)$ replaced by $m(\xi/8s_\nu)$.

The decomposition (15-2) can be modified so that F_ν^0, F_ν^∞ and $F_\nu^\flat$ remain real-valued and even, without sacrificing any of its desired properties. First replace each summand by its real part. Then replace $F_\nu^0(x)$ by $\frac{1}{2}F_\nu^0(x) + \frac{1}{2}F_\nu^0(-x)$, and similarly for F_ν^∞ and $F_\nu^\flat$.

The remainder $(1-\zeta)g_{\nu,j}^\infty$, which is neglected in the construction of F_ν^∞, gives rise to one of several summands which contribute to $F_\nu^\flat$. Because $S_\nu \to \infty$, and because the cutoff function m is smooth and compactly supported, $\|(1-\zeta)g_{\nu,j}^\infty\|_{L^2(\mathbb{R}^2)} \to 0$ as $\nu \to \infty$.

From the fact that $s_\nu \to \infty$ and the relation $S_\nu \ge s_\nu^3$, it follows easily that $\langle F_\nu^0, F_\nu^\infty\rangle \to 0$ as $\nu \to \infty$. Therefore since $\|F_\nu^\flat\|_2 \to 0$,

$$\|f_\nu\|_2^2 - \|F_\nu^0\|_2^2 - \|F_\nu^\infty\|_2^2 \to 0$$
$$\|F_\nu^0\|_2^2 + \|F_\nu^\infty\|_2^2 \to 1 = \|f_\nu\|_2^2.$$

As in Section 14, the relation $S_\nu \ge s_\nu^3 \to \infty$ also leads to

$$\|F_\nu^0\sigma * F_\nu^\infty\sigma\|_{L^2(\mathbb{R}^3)} \to 0. \tag{15-3}$$

This requires several substeps. These are entirely parallel to those in Section 12 and Section 14, so the details are omitted.

We need to know that

$$\liminf_{\nu\to\infty} \|F_\nu^\infty\|_2 > 0.$$

This is less apparent than was the analogous statement in the proof of Proposition 2.15, because F_ν^∞ is defined here as a sum over j that recombines different terms resulting from our partition of unity, and it must be shown that this summation does not introduce unwanted cancellation. Indeed, suppose to the contrary that $\|F_\nu^\infty\|_2 \to 0$ for a subsequence of values of ν. Then there must exist an index i such that for a certain sub-subsequence,

$$\int_{|\xi|\ge S_\nu} |\widehat{g_{\nu,i}}(\xi)|^2 \, d\xi \gtrsim 1. \tag{15-4}$$

Pass to such a sub-subsequence, substitute the representation $f_\nu = F_\nu^0 + F_\nu^\infty + F_\nu^\flat$ into the definition $g_{\nu,i} = \phi_i^*(\eta_i f_\nu)$, and consider $\widehat{g_{\nu,i}}(\xi)$ for $|\xi| \gtrsim S_\nu$. The contribution of F_ν^0 to this Fourier transform in this regime tends to zero in $L^2(d\xi)$ norm, because $S_\nu/s_\nu \to \infty$. The contribution of $F_\nu^\flat$ tends to zero in L^2 norm, because $F_\nu^\flat$ itself does so. Therefore the contribution of F_ν^∞ to the integral (15-4) cannot tend to zero. Therefore $\|F_\nu^\infty\|_{L^2(\mathbb{S}^2)}$ cannot tend to zero.

Since the $L^2(\mathbb{S}^2)$ norms of both F_ν^∞ and F_ν^0 enjoy strictly positive lower bounds, the small cross-term bound (15-3) implies as in the proof of Lemma 12.2 that

$$\limsup_{\nu\to\infty} \|f_\nu\sigma * f_\nu\sigma\|_2^2 < \mathbf{S}^4,$$

contradicting the assumption that $\{f_\nu\}$ is an extremizing sequence.

16. Constants are local maxima

Theorem 1.8 asserts that constant functions are local maxima. Define

$$\Psi(f) = \|f\sigma * f\sigma\|_{L^2(\mathbb{R}^3)}^2 \quad \text{and} \quad \Phi(f) = \frac{\Psi(f)}{\|f\|_{L^2(\mathbb{S}^2)}^4}.$$

Denote by $\mathbf{1}$ the constant function $\mathbf{1}(x) = 1$ for all $x \in \mathbb{S}^2$.

Proof of Theorem 1.8. Since $\Phi(f) = \Phi(tf)$ for all $t > 0$, and since $\Phi(f) \le \Phi(|f|)$, we may restrict attention to functions of the form $f = \mathbf{1} + \varepsilon g$ where $0 \le \varepsilon \le \delta$, $g \perp \mathbf{1}$, g is real-valued, and $\|g\|_{L^2(\mathbb{S}^2)} = 1$. We may further assume that $g(-x) = g(x)$, by Proposition 2.7.

The constant function $\mathbf{1}$ is a critical point for Φ. Indeed, by rotation symmetry, $f = \mathbf{1}$ satisfies the generalized Euler–Lagrange equation $f = \lambda(f\sigma * f\sigma * \tilde{f}\sigma)\big|_{\mathbb{S}^2}$ that characterizes critical points.

A straightforward calculation gives the Taylor expansion

$$\Phi(\mathbf{1}+\varepsilon g) = \Phi(\mathbf{1}) + \varepsilon^2 \|\mathbf{1}\|_{L^2(\mathbb{S}^2)}^{-4} \big(6\langle g\sigma * g\sigma,\ \sigma * \sigma\rangle - 2\Psi(\mathbf{1})\|\mathbf{1}\|_2^{-2}\|g\|_2^2\big) + O(\varepsilon^3),$$

where $O(\varepsilon^3)$ denotes a quantity whose absolute value is majorized by $C\varepsilon^3$, uniformly for $g \in L^2(\mathbb{S}^2)$ satisfying $\|g\|_2 \le 1$. Thus it suffices to show that

$$\sup_{\|g\|_2=1} 6\langle g\sigma * g\sigma,\ \sigma * \sigma\rangle < 2\Psi(\mathbf{1})\|\mathbf{1}\|_2^{-2}.$$

The quantities $\Psi(\mathbf{1})$ and $\|\mathbf{1}\|_2$ can be evaluated explicitly. Firstly, $\|\mathbf{1}\|_2^2 = \sigma(\mathbb{S}^2) = 4\pi$. Secondly,

$$(\sigma * \sigma)(x) = 2\pi |x|^{-1} \chi_{|x|\le 2}.$$

Indeed, it follows from trigonometry that $\sigma * \sigma(x) = a|x|^{-1}\chi_{|x|\le 2}$ for some $a > 0$, and a can be evaluated by

$$(4\pi)^2 = \sigma(\mathbb{S}^2)^2 = \int_{\mathbb{R}^3} (\sigma * \sigma)(x)\, dx = \int_0^2 ar^{-1} \cdot 4\pi r^2\, dr = 8\pi a.$$

Therefore

$$\Psi(\mathbf{1}) = \int_{\mathbb{R}^3} \big(\sigma * \sigma(x)\big)^2\, dx = \int_{\mathbb{R}^3} 4\pi^2 |x|^{-2}\, dx = 4\pi^2 \int_0^2 r^{-2} \cdot 4\pi r^2\, dr = 4\pi^2 \cdot 4\pi \cdot 2 = 32\pi^3.$$

Therefore it suffices to prove that

$$\sup_{\|g\|_2 = 1} \langle g\sigma * g\sigma,\ \sigma * \sigma \rangle < \tfrac{1}{3} \cdot 32\pi^3 \cdot (4\pi)^{-1} = \tfrac{8}{3}\pi^2,$$

where the supremum is taken over all real-valued, even $g \in L^2(\mathbb{S}^2)$ satisfying $\|g\|_2 = 1$ and $\int g\, d\sigma = 0$.

The following key bound will be established below.

Lemma 16.1. *For all real-valued even functions $g \in L^2(\mathbb{S}^2)$ satisfying $\int g\, d\sigma = 0$,*

$$\Big| \iint_{\mathbb{S}^2 \times \mathbb{S}^2} g(x) g(y) |x - y|^{-1}\, d\sigma(x)\, d\sigma(y) \Big| \le \tfrac{4}{5}\pi \|g\|^2_{L^2(\mathbb{S}^2)}.$$

The factor $\frac{4}{5}\pi$ is optimal, and is attained if and only if g is a spherical harmonic of degree 2.

Now for such g satisfying $\|g\|_2 = 1$,

$$\begin{aligned} \langle g\sigma * g\sigma,\ \sigma * \sigma \rangle &= \langle g\sigma * (\sigma * \sigma),\ g \rangle \\ &= 2\pi \iint_{\mathbb{S}^2 \times \mathbb{S}^2} g(x) g(y) |x - y|^{-1}\, d\sigma(x)\, d\sigma(y) \le 2\pi \cdot \tfrac{4}{5}\pi = \tfrac{8}{5}\pi^2 < \tfrac{8}{3}\pi^2, \end{aligned}$$

completing the proof of Theorem 1.8. □

Proof of Lemma 16.1. We first recall the Funk–Hecke formula in the theory of spherical harmonics, see e.g., [Müller 1998, page 29] or [Xu 2000, Theorem A].

Theorem 16.2 (Funk–Hecke formula). *Let $d \ge 2$ and $k \ge 0$ be integers. Let f be a continuous function on $[-1, 1]$ and Y_k be a spherical harmonic of degree k, on the sphere S^d. Then for any $x \in S^d$,*

$$\int_{S^d} f(x \cdot y) Y_k(y) d\sigma(y) = \lambda_k Y_k(x),$$

where $x \cdot y$ is the usual inner product in $\mathbb{R}^{d+1}$, and

$$\lambda_k = \frac{\omega_d \int_{-1}^1 f(t) C_k^{(d-1)/2}(t)(1 - t^2)^{(d-2)/2} dt}{C_k^{(d-1)/2}(1) \int_{-1}^1 (1 - t^2)^{(d-2)/2} dt},$$

where $\omega_d := 2\pi^{(d+1)/2}/\Gamma((d+1)/2)$ *denotes the surface area of the unit sphere* S^d *and* $C_k^\nu(t)$ *is the Gegenbauer polynomial defined by the generating function*

$$(1-2rt+r^2)^{-\nu} = \sum_{k=0}^{\infty} C_k^\nu r^k \tag{16-1}$$

for $0 \le r < 1$ *and* $-1 \le t \le 1$ *and* $\nu > 0$.

For $\nu = 1/2$ and $t = 1$, the generating formula becomes $(1-r)^{-2/2} = \sum_{k=0}^{\infty} C_k^{1/2} r^k$, so

$$C_k^{1/2} = 1 \quad \text{for all } k \ge 0.$$

For $d = 2$, $(d-2)/2 = 0$ and $\omega_d = 4\pi$, and the relevant index ν is $\nu = (d-1)/2 = 1/2$. Therefore for $d = 2$,

$$\lambda_k = 2\pi \int_{-1}^{1} f(t) C_k^{1/2}(t)\,dt. \tag{16-2}$$

Choosing $\nu = 1/2$ and setting $r = 1$ in the generating function (16-1), we obtain

$$(2-2t)^{-1/2} = \sum_{k=0}^{\infty} C_k^{1/2}(t).$$

This formula is not entirely valid, since (16-1) only applies for $r < 1$; but all calculations below can be justified by writing the corresponding formulas for $r < 1$ and then passing to the limit $r = 1$. We will omit these details, and work directly with $r = 1$.

We also recall the following fact in [Stein and Weiss 1971, Chapter 4, Corollary 2.16]: For $\mathbb{S}^2$, the polynomials $C_k^{1/2}(t)$ for $k = 0, 1, \ldots$ are mutually orthogonal with respect to the inner product $\langle f, g\rangle = \int_{-1}^{1} f(t) g(t)\,dt$. So for $f = (2-2t)^{-1/2}$ in (16-2) and for any $k \ge 0$, by orthogonality,

$$\begin{aligned}\lambda_k = 2\pi \int_{-1}^{1} (2-2t)^{-1/2} C_k^{1/2}(t)\,dt &= 2\pi \int_{-1}^{1} \sum_{m=0}^{\infty} C_m^{1/2}(t) C_k^{1/2}(t)\,dt \\ &= 2\pi \int_{-1}^{1} (C_k^{1/2}(t))^2\,dt = \frac{4\pi}{2k+1},\end{aligned}$$

where the last identity follows from the normalized value of $C_k^{1/2}(t)$ over $(-1, 1)$, see e.g., [Andrews et al. 1999, page 461] or [Müller 1998, 10.15, page 54]. Hence for $f(t) = (2-2t)^{-1/2}$, for $x \in \mathbb{S}^2$,

$$\int_{\mathbb{S}^2} f(x\cdot y) Y_k(y)\,d\sigma(y) = \frac{4\pi}{2k+1} Y_k(x) \quad \text{for all } k \ge 0.$$

Now return to $\iint g(x)g(y)|x-y|^{-1}\,d\sigma(x)\,d\sigma(y)$. Here $|x-y|^{-1} = (2-2x\cdot y)^{-1/2} = f(x\cdot y)$, where $f(t) = (2-2t)^{-1/2}$. Since all spherical harmonics of odd degrees are odd, and since $g \perp \mathbf{1}$, g may be expanded as $g = \sum_{k=1}^{\infty} Y_{2k}$, where each Y_{2k} is a spherical harmonic of degree $2k$. These are of course pairwise orthogonal in $L^2(\mathbb{S}^2)$. Therefore

$$\iint g(x)g(y)|x-y|^{-1}\,d\sigma(x)\,d\sigma(y) = \sum_{k=1}^{\infty}\langle \lambda_{2k}Y_{2k}, Y_{2k}\rangle$$
$$= \sum_{k=1}^{\infty}\Big\langle \frac{4\pi}{2(2k)+1}Y_{2k}, Y_{2k}\Big\rangle \le \frac{4\pi}{5}\sum_{k=1}^{\infty}\|Y_{2k}\|_2^2 = \frac{4\pi}{5}\|g\|_2^2.$$

This completes the proof of Lemma 16.1. □

Remark 16.3. Consider inequalities of the modified form

$$\int_{\mathbb{R}^3}\big|(f\sigma * f\sigma)(x)\big|^2\,w(x)\,dx \le C\|f\|^4_{L^4(\mathbb{S}^2)}, \tag{16-3}$$

where $w \ge 0$ is any radial weight. The modification consists in placing the L^4 norm on the right side of the inequality instead of the L^2 norm.

If the inequality holds for some $C < \infty$, and if w satisfies $|\lambda_k(w)| \le \lambda_0(w)$, where

$$\lambda_k(w) = 2\pi\int_{-1}^{1} w((2+2t)^{1/2})(2+2t)^{-1/2}\,C_k^{1/2}(t)\,dt,$$

then constant functions are (global) extremals. This holds in particular for $w \equiv 1$.

This is proved as follows, in the spirit of Foschi [2007]. We may assume that $f \ge 0$.

$$\int_{\mathbb{R}^3}(f\sigma * f\sigma)(x)^2 w(x)\,dx \le \int_{\mathbb{R}^3}\big((f^2\sigma*\sigma)(x)\big)^2 w(x)\,dx$$
$$= 2\pi\iint_{\mathbb{S}^2\times\mathbb{S}^2} f^2(x)f^2(y)|x+y|^{-1}w(|x+y|)\,d\sigma(x)\,d\sigma(y).$$

The first inequality follows from the Cauchy–Schwarz inequality, and is an equality if f is constant modulo null sets on almost every circle (that is, the intersection of $\mathbb{S}^2$ with an affine plane) in $\mathbb{S}^2$; thus if and only if f is constant modulo σ-null sets. Expand $f^2 = \sum_{k=0}^{\infty} Y_k$ in spherical harmonics. Then

$$2\pi\iint_{\mathbb{S}^2\times\mathbb{S}^2} f^2(x)f^2(y)|x+y|^{-1}w(|x+y|)\,d\sigma(x)\,d\sigma(y) = 2\pi\sum_{k=0}^{\infty}\lambda_k\|Y_k\|_2^2 \le 2\pi\sup_k\lambda_k\|f\|_4^4,$$

for certain coefficients λ_k that depend only on w. If there is a valid inequality (16-3) with $C < \infty$, then $\lambda_0 < \infty$. Thus constant functions are extremizers. If $\max_{k\ne 0}|\lambda_k(w)| < \lambda_0(w)$, then f is an extremizer if and only if f^2 has a spherical harmonic expansion with $Y_k = 0$ for all $k \ge 1$, that is, if and only if f^2 is constant. For $f \ge 0$, this forces f to be constant. □

17. A variational calculation

Recall the notation $e_\xi(x) = e^{x\cdot\xi}$. It is natural to study $\|\widehat{f\sigma}\|_4/\|f\|_2$ for $f(x) = e_\xi(x)$, for several reasons.

(i) Extremizers for the paraboloid $\mathbb{P}^2 = \{x : x_3 = \frac{1}{2}|x'|^2\}$, where $x' = (x_1, x_2)$, are Gaussian functions of x'; but these are simply restrictions to $\mathbb{P}^2$ of simple exponentials $e^{x\cdot\xi}$ for $\xi \in \mathbb{C}^3$ satisfying $\mathrm{Re}(\xi_3) < 0$.

(ii) $(f\sigma * f\sigma)(x)$ is expressed for each x as an integral of a product of two factors. When $f = e_\xi$, the integrand becomes a constant for each x, and hence the Cauchy–Schwarz inequality becomes an equality when applied to each such integral in an appropriate way. Such equalities are the key to one proof [Foschi 2007] that Gaussians are extremal for $\mathbb{P}^2$.

(iii) The functional $\|e_\xi\sigma * e_\xi\sigma\|_2/\|e_\xi\|_2^2$ is susceptible to a perturbative analysis for large $|\xi|$.

(iv) This analysis appears more likely to be generalizable to other manifolds than $\mathbb{S}^2$ than does the calculation of Lemma 2.4 for $f \equiv 1$.

For these reasons, we carry out in this section a perturbative analysis of $\|e_\xi\sigma * e_\xi\sigma\|_2/\|e_\xi\|_2^2$, thereby establishing Proposition 2.19.

We will work with functions concentrated principally in a very small neighborhood of the north pole $(0, 0, 1)$. A point $z \approx (0, 0, 1)$ in $\mathbb{S}^2$ can be written as

$$(y, (1-|y|^2)^{1/2}) = (y, 1 - \tfrac{1}{2}|y|^2 - \tfrac{1}{8}|y|^4 + O(|y|^6)),$$

where $y \in \mathbb{R}^2$ and $|y| < 1$. Let σ denote surface measure on $\mathbb{S}^2$;

$$d\sigma = (1 + \tfrac{1}{2}|y|^2 + O(|y|^4))\,dy.$$

For $z \in \mathbb{S}^2$ and $\varepsilon > 0$ define

$$f_\varepsilon(z) = \varepsilon^{-1/2} e^{(z_3-1)/\varepsilon} \chi_{|(z_1,z_2)|<\frac{1}{2}} \chi_{z_3>0}.$$

Within the domain of f_ε, the mapping $(z_1, z_2, z_3) \leftrightarrow (z_1, z_2)$ is a one-to-one correspondence between $\mathbb{S}^2$ and a ball in $\mathbb{R}^2$.

We observe that f_ε is essentially $\varepsilon^{-1/2}e^{-1/\varepsilon}e_\xi$, where $\xi = (0, 0, \varepsilon^{-1})$; the two functions differ by $O(e^{-c/\varepsilon})$ in L^2 norm for some $c > 0$. The cutoff functions are inserted for convenience in the calculation.

For $(t, x) \in \mathbb{R}^{1+2}$, define

$$u_\varepsilon(t, x) = \int_{\mathbb{S}^2} f_\varepsilon(z) e^{-i(x,t)\cdot z}\,d\sigma(z),$$

where of course $(x, t)\cdot z = x_1z_1 + x_2z_2 + tz_3$. Then

$$\begin{aligned} u_\varepsilon(t, x) &= \varepsilon^{-1/2} \int_{\mathbb{S}^2} e^{(z_3-1)/\varepsilon} e^{-ix\cdot(z_1,z_2)} e^{-itz_3}\,\tilde{\chi}(z)\,d\sigma(z) \\ &= \varepsilon^{-1/2} e^{-it} \int_{\mathbb{R}^2} e^{(-|y|^2/2-|y|^4/8+O(|y|^6))\varepsilon^{-1}} \\ &\qquad\qquad \cdot e^{-ix\cdot y} e^{-it(-|y|^2/2-|y|^4/8+O(|y|^6))} (1 + |y|^2/2 + O(|y|^4))\chi(y)\,dy, \end{aligned}$$

where $\tilde{\chi}$ and χ denote disks centered respectively at $(0,0,1)\in\mathbb{S}^2$ and $0\in\mathbb{R}^2$, which are independent of ε. A change of variables gives

$$u_\varepsilon(t,x)=\varepsilon^{1/2}e^{-it}\int_{\mathbb{R}^2}e^{-i\varepsilon^{1/2}x\cdot y}e^{-(1-i\varepsilon t)(|y|^2/2+\varepsilon|y|^4/8+O(\varepsilon^{-1}|\varepsilon^{1/2}y|^6))}\cdot(1+\varepsilon|y|^2/2+O(|\varepsilon^{1/2}y|^4))\chi(\varepsilon^{1/2}y)\,dy.$$

Setting

$$\begin{aligned}v_\varepsilon(t,x)&=e^{-it/\varepsilon}\varepsilon^{-1/2}u_\varepsilon(-\varepsilon^{-1}t,\varepsilon^{-1/2}x)\\&=\int_{\mathbb{R}^2}e^{-ix\cdot y}e^{-(1+it)(|y|^2/2+\varepsilon|y|^4/8+O(\varepsilon^{-1}|\varepsilon^{1/2}y|^6)}(1+\varepsilon|y|^2/2+O(|\varepsilon^{1/2}y|^4))\chi(\varepsilon^{1/2}y)\,dy,\end{aligned}$$

we have

$$\|v_\varepsilon\|^4_{L^4(\mathbb{R}^3)}=\|u_\varepsilon\|^4_{L^4(\mathbb{R}^3)}. \tag{17-1}$$

Set

$$w_\varepsilon(t,x)=\int_{\mathbb{R}^2}e^{-ix\cdot y}e^{-(1+it)(|y|^2/2+\varepsilon|y|^4/8)}(1+\tfrac{1}{2}\varepsilon|y|^2)\,dy\quad\text{for }\varepsilon\ge 0.$$

Using the exact definition of f_ε rather than the approximate expressions above, it is routine to verify that

$$\|w_\varepsilon\|_4^4=\|v_\varepsilon\|_4^4+O(\varepsilon^2)\quad\text{as }\varepsilon\to 0^+.$$

Since we are interested in first variations with respect to ε of the L^4 norm at $\varepsilon=0$, it will suffice to analyze $\|w_\varepsilon\|_4^4$. Also introduce

$$g_\varepsilon(y)=e^{-|y|^2/2-\varepsilon|y|^4/8}\quad\text{and}\quad d\sigma_\varepsilon(y)=(1+\varepsilon|y|^2/2)\,dy.$$

Then

$$\|f_\varepsilon\|^2_{L^2(\sigma)}=\|g_\varepsilon\|^2_{L^2(\sigma_\varepsilon)}+O(\varepsilon^2).$$

Although f_ε is not well defined in the limit $\varepsilon=0$, the limit $\lim_{\varepsilon\to0^+}\|f_\varepsilon\|_2^2>0$ does exist, and we will abuse notation by writing $\|f_0\|_2^2$ to denote this quantity. We have

$$\|f_0\|_2^2=\int_{\mathbb{R}^2}e^{-|y|^2}\,dy.$$

It is a routine exercise to verify that $\varepsilon\mapsto\|v_\varepsilon\|_4^4$ is a C^∞ function on $[0,\infty)$; hence the same goes for $\|w_\varepsilon\|_4^4$, and for $\|u_\varepsilon\|_4^4$ by (17-1). Similarly, $\varepsilon\mapsto\|f_\varepsilon\|_2^2$ is C^∞ on $[0,\infty)$.

Consider the functional

$$\Psi(\varepsilon)=\log\frac{\|u_\varepsilon\|^4_{L^4}}{\|f_\varepsilon\|^4_{L^2}},$$

which is initially defined for $\varepsilon>0$ but extends continuously and differentiably to $\varepsilon=0$. Its derivative is

$$\partial_\varepsilon\big|_{\varepsilon=0}\Psi(\varepsilon)=\frac{\partial_\varepsilon\|w_\varepsilon\|_4^4\big|_{\varepsilon=0}}{\|w_0\|_4^4}-2\frac{\partial_\varepsilon\big|_{\varepsilon=0}\|g_\varepsilon\|_2^2}{\|g_0\|_2^2}, \tag{17-2}$$

and of course $\Psi(0) = \log(\mathscr{R}_{\mathbb{P}^2}^4)$, where $\mathscr{R}_{\mathbb{P}^2}$ from (1-3) is the optimal constant for the adjoint restriction inequality for the paraboloid.

Lemma 17.1.
$$\frac{\partial \Psi}{\partial \varepsilon}\Big|_{\varepsilon=0} > 0.$$

Proposition 2.19 follows, since by radial symmetry, $\|e_\xi\sigma * e_\xi\sigma\|_2/\|e_\xi\|_2^2$ depends only on $|\xi|$.

The most involved calculation is that of the numerator in the first term of (17-2). To begin that calculation,

$$\begin{aligned}\partial_\varepsilon\big|_{\varepsilon=0} w_\varepsilon(t,x) &= \int \big(-\tfrac{1}{8}(1+it)|y|^4 + \tfrac{1}{2}|y|^2\big) e^{-ix\cdot y} e^{-(1+it)|y|^2/2}\,dy \\ &= \big(-\tfrac{1}{8}(1+it)(-i/2)^{-2}\partial_t^2 + \tfrac{1}{2}(-i/2)^{-1}\partial_t\big) \int e^{-ix\cdot y} e^{-(1+it)|y|^2/2}\,dy \\ &= \big(\tfrac{1}{2}(1+it)\partial_t^2 + i\partial_t\big) \int e^{-ix\cdot y} e^{-(1+it)|y|^2/2}\,dy \\ &= \big(\tfrac{1}{2}(1+it)\partial_t^2 + i\partial_t\big) w_0(t,x) \\ &= \big(\tfrac{1}{2}(1+it)\partial_t^2 + i\partial_t\big) c_0(1+it)^{-1} e^{-|x|^2/2(1+it)},\end{aligned}$$

where c_0 is a positive constant whose precise value will play no role, since it will ultimately appear in both the numerator and denominator of a certain ratio.

Define
$$\phi(t,x) = -\tfrac{1}{2}|x|^2(1+it)^{-1} - \log(1+it),$$

so that $w_0 = c_0 e^\phi$. The last quantity above may be written as

$$\begin{aligned}c_0\big(\tfrac{1}{2}(1+it)\partial_t^2 + i\partial_t\big)e^\phi &= \tfrac{1}{2}c_0(1+it)(\phi_t^2+\phi_{tt})e^\phi + c_0 i\phi_t e^\phi \\ &= \big(\tfrac{1}{2}(1+it)(\phi_t^2+\phi_{tt}) + i\phi_t\big)w_0,\end{aligned}$$

where ϕ_t and ϕ_{tt} denote respectively the first and second partial derivatives of ϕ with respect to t.

Now
$$\begin{aligned}\phi_t &= \tfrac{i}{2}|x|^2(1+it)^{-2} - i(1+it)^{-1}, \\ \phi_{tt} &= \tfrac{i}{2}(-2i)|x|^2(1+it)^{-3} - i(-i)(1+it)^{-2} = |x|^2(1+it)^{-3} - (1+it)^{-2}, \\ \phi_t^2 &= -\tfrac{1}{4}|x|^4(1+it)^{-4} + |x|^2(1+it)^{-3} - (1+it)^{-2},\end{aligned}$$

so
$$\phi_t^2 + \phi_{tt} = -\tfrac{1}{4}|x|^4(1+it)^{-4} + 2|x|^2(1+it)^{-3} - 2(1+it)^{-2}.$$

Consequently

$$\begin{aligned}&\tfrac{1}{2}(1+it)(\phi_t^2+\phi_{tt}) + i\phi_t \\ &\qquad = -\tfrac{1}{8}|x|^4(1+it)^{-3} + |x|^2(1+it)^{-2} - (1+it)^{-1} - \tfrac{1}{2}|x|^2(1+it)^{-2} + (1+it)^{-1} \\ &\qquad = -\tfrac{1}{8}|x|^4(1+t^2)^{-3}(1-it)^3 + \tfrac{1}{2}|x|^2(1+t^2)^{-2}(1-it)^2,\end{aligned}$$

whose real part is

$$\operatorname{Re}\big(\tfrac{1}{2}(1+it)\big(\phi_t^2+\phi_{tt}\big)+i\phi_t\big) = -\tfrac{1}{8}|x|^4(1+t^2)^{-3}(1-3t^2)+\tfrac{1}{2}|x|^2(1+t^2)^{-2}(1-t^2).$$

Now $\partial_\varepsilon \|w_\varepsilon\|_4^4 = 4\int |w_\varepsilon|^4 \operatorname{Re}(\partial_\varepsilon w_\varepsilon / w_\varepsilon)$, and therefore

$$\begin{aligned}
\partial_\varepsilon \|w_\varepsilon\|_4^4\big|_{\varepsilon=0} &= 4\iint \operatorname{Re}\big(\tfrac{1}{2}(1+it)(\phi_t^2+\phi_{tt})+i\phi_t\big)|w_0(t,x)|^4\,dx\,dt\\
&= c_0^4\int_{\mathbb{R}}\int_{\mathbb{R}^2}\big(-\tfrac{1}{2}|x|^4(1+t^2)^{-3}(1-3t^2)+2|x|^2(1+t^2)^{-2}(1-t^2)\big)\\
&\qquad\qquad\cdot(1+t^2)^{-2}|e^{-|x|^2/2(1+it)}|^4\,dx\,dt\\
&= c_0^4\int_{\mathbb{R}}\int_{\mathbb{R}^2}\big(-\tfrac{1}{2}|x|^4(1+t^2)^{-3}(1-3t^2)+2|x|^2(1+t^2)^{-2}(1-t^2)\big)\\
&\qquad\qquad\cdot(1+t^2)^{-2}\,e^{-2|x|^2/(1+t^2)}\,dx\,dt.
\end{aligned}$$

Substituting $x=(1+t^2)^{1/2}\tilde{x}$ and then replacing $\tilde{x}$ by x gives

$$\partial_\varepsilon \|w_\varepsilon\|_4^4\big|_{\varepsilon=0} = c_0^4\int_{\mathbb{R}}\int_{\mathbb{R}^2}\big(-\tfrac{1}{2}|x|^4(1-3t^2)+2|x|^2(1-t^2)\big)(1+t^2)^{-2}e^{-2|x|^2}\,dx\,dt.$$

By substituting $x=2^{-1/2}y$ in $\mathbb{R}^2$ and then $r=s^{1/2}$ in $(0,\infty)$, we derive the identities

$$\begin{aligned}
\int_{\mathbb{R}^2} e^{-2|x|^2}\,dx &= \tfrac{1}{2}\int_{\mathbb{R}^2} e^{-|y|^2}\,dy = \pi\int_0^\infty e^{-r^2}\,r\,dr = \tfrac{1}{2}\pi\int_0^\infty e^{-s}\,ds = \frac{\pi}{2},\\
\int_{\mathbb{R}^2}|x|^2e^{-2|x|^2}\,dx &= \frac{\pi}{4}\int_0^\infty se^{-s}\,ds = \frac{\pi}{4},\\
\int_{\mathbb{R}^2}|x|^4e^{-2|x|^2}\,dx &= \frac{\pi}{8}\int_0^\infty s^2e^{-s}\,ds = \frac{\pi}{4}.
\end{aligned}$$

Recall also that

$$\int_{\mathbb{R}}(1+t^2)^{-1}\,dt = \pi \quad\text{and}\quad \int_{\mathbb{R}}(1+t^2)^{-2}\,dt = \frac{\pi}{2}.$$

Using these formulas we obtain

$$\begin{aligned}
\partial_\varepsilon \|w_\varepsilon\|_4^4\big|_{\varepsilon=0} &= c_0^4\int_{\mathbb{R}}\Big(-\tfrac{1}{2}(1-3t^2)\frac{\pi}{4}+2(1-t^2)\frac{\pi}{4}\Big)(1+t^2)^{-2}\,dt\\
&= \tfrac{\pi}{4}c_0^4\int_{\mathbb{R}}(-\tfrac{1}{2}t^2+\tfrac{3}{2})(1+t^2)^{-2}\,dt\\
&= \tfrac{\pi}{4}c_0^4\int_{\mathbb{R}}\big(-\tfrac{1}{2}(1+t^2)^{-1}+2(1+t^2)^{-2}\big)\,dt = \tfrac{\pi}{4}c_0^4\Big(-\frac{\pi}{2}+2\frac{\pi}{2}\Big) = c_0^4\frac{\pi^2}{8}.
\end{aligned}$$

On the other hand,

$$\|w_0\|_4^4 = c_0^4\int_{\mathbb{R}}\int_{\mathbb{R}^2}(1+t^2)^{-2}e^{-2|x|^2/(1+t^2)}\,dx\,dt = c_0^4\int_{\mathbb{R}}\int_{\mathbb{R}^2}(1+t^2)^{-1}e^{-2|y|^2}\,dy\,dt = c_0^4\tfrac{1}{2}\pi^2.$$

Therefore

$$\frac{\partial_\varepsilon \|w_\varepsilon\|_4^4\big|_{\varepsilon=0}}{\|w_0\|_4^4} = \frac{\pi^2c_0^4/8}{\pi^2c_0^4/2} = \frac{1}{4}.$$

The variation of $\|g_\varepsilon\|_2^2$ must also be taken into account:

$$\partial_\varepsilon \int_{\mathbb{R}^2} g_\varepsilon(y)^2 \, d\sigma_\varepsilon(y) \Big|_{\varepsilon=0} = \partial_\varepsilon \int_{\mathbb{R}^2} e^{-|y|^2-\varepsilon\frac{1}{4}|y|^4}(1+\varepsilon\tfrac{1}{2}|y|^2)\, dy \Big|_{\varepsilon=0}$$
$$= \int_{\mathbb{R}^2} (-\tfrac{1}{4}|y|^4 + \tfrac{1}{2}|y|^2) e^{-|y|^2}\, dy = -\frac{2\pi}{4} + \frac{\pi}{2} = 0.$$

Therefore $2\partial_\varepsilon \|g_\varepsilon\|^2_{L^2(\sigma_\varepsilon)}\big|_{\varepsilon=0} \big/ \|g_0\|_2^2 = 0$. Putting it all together, $\partial_\varepsilon \Psi(\varepsilon)\big|_{\varepsilon=0} = \frac{1}{4} - 0 > 0$.

18. Proof of Lemma 6.1

Proof of Lemma 6.1. Suppose that $f = \chi_E$ is the characteristic function of a set E. We will begin by showing that there exist $C < \infty$ and exponents $s, t > 0$ such that for any set E and any index k,

$$\sum_j |\mathscr{C}_k^j|^2 \Big(|\mathscr{C}_k^j|^{-1} \int_{\mathscr{C}_k^j} |\chi_E|^p \Big)^{4/p} \le C|E|^2 \cdot \min\big(2^{-2k}|E|^{-1}, 2^{2k}|E|\big)^t \cdot \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|E| + |\mathscr{C}_k^i|} \Big)^s. \tag{18-1}$$

Indeed,

$$\begin{aligned}\sum_j |\mathscr{C}_k^j|^2 \Big(|\mathscr{C}_k^j|^{-1} \int_{\mathscr{C}_k^j} \chi_E^p \Big)^{4/p} &= \sum_j |\mathscr{C}_k^j|^2 |E \cap \mathscr{C}_k^j|^{4/p} |\mathscr{C}_k^j|^{-4/p} \\ &\le \sum_j |E \cap \mathscr{C}_k^j| \cdot \max_i \big(|E \cap \mathscr{C}_k^i|^{4/p-1} |\mathscr{C}_k^i|^{2-4/p} \big) \\ &= |E| \max_i \big(|E \cap \mathscr{C}_k^i|^{4/p-1} |\mathscr{C}_k^i|^{2-4/p} \big).\end{aligned}$$

The analysis now splits into two cases. Note that $|\mathscr{C}_k^j| \sim 2^{-2k}$ uniformly for all indices j and k. If $2^{-2k} \ge |E|$, then

$$\begin{aligned}|E| \max_i \big(|E \cap \mathscr{C}_k^i|^{4/p-1} |\mathscr{C}_k^i|^{2-4/p} \big) &\le |E|^2 \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|\mathscr{C}_k^i|} \Big)^{4/p-2} \\ &\le |E|^2 (2^{2k}|E|)^{2/p-1} \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|\mathscr{C}_k^i|} \Big)^{2/p-1}.\end{aligned}$$

Since $1 \le p < 2$, we have $2/p - 1 > 0$ and hence this is a bound of the required form (18-1). If instead $2^{-2k} < |E|$, then since $4/p - 1 > 1 \ge \frac{1}{2}$,

$$\begin{aligned}|E| \max_i \big(|E \cap \mathscr{C}_k^i|^{4/p-1} |\mathscr{C}_k^i|^{2-4/p} \big) &= |E|^2 (2^{2k}|E|)^{-1} \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|\mathscr{C}_k^i|} \Big)^{4/p-1} \\ &\le |E|^2 (2^{2k}|E|)^{-1} \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|\mathscr{C}_k^i|} \Big)^{1/2} \\ &= |E|^2 (2^{2k}|E|)^{-1/2} \max_i \Big(\frac{|E \cap \mathscr{C}_k^i|}{|E|} \Big)^{1/2},\end{aligned}$$

which again is a bound of the desired form. Thus (18-1) is proved.

Next consider a general function $f \in L^2(\mathbb{S}^2)$. By sacrificing a constant factor in the inequality, we may assume that f takes the form $f = \sum_{\alpha=-\infty}^{\infty} 2^\alpha \chi_{E_\alpha}$, where the sets E_α are pairwise disjoint and $|E_\alpha| < \infty$. Invoking the preceding analysis for each summand together with the triangle inequality for the sum with respect to α yields

$$\|f\|_{X_p}^4 \le C \sum_k \Big(\sum_\alpha 2^\alpha |E_\alpha|^{1/2} \cdot \min\big(2^{-2k}|E_\alpha|^{-1}, 2^{2k}|E_\alpha|\big)^{t/4} \cdot \max_i \Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \Big)^{s/4} \Big)^4 \tag{18-2}$$

$$\le C \Big(\sum_\alpha 2^{4\alpha} |E_\alpha|^2 \max_{k,i} \Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \Big)^s \Big)^{1/2} \|f\|_2^2. \tag{18-3}$$

The second inequality in (18-3) is deduced as follows. For each integer r define

$$a_r = \sum_{\beta:|E_\beta| \in [2^r, 2^{r+1})} 2^\beta |E_\beta|^{1/2} \max_{m,i} \Big(\frac{|E_\beta \cap \mathscr{C}_m^i|}{|E_\beta| + |\mathscr{C}_m^i|} \Big)^{s/4} \quad \text{and} \quad b_{k,r} = \min\big(2^{-(r+2k)t/4}, 2^{(r+2k)t/4}\big).$$

Then by (18-2),

$$\|f\|_{X_p} \le C \Big(\sum_{k=0}^{\infty} \big(\sum_{r=-\infty}^{\infty} a_r b_{k,r} \big)^4 \Big)^{1/4}$$

$$\le C \Big(\sum_{k=0}^{\infty} \big(\sum_r a_r^4 b_{k,r} \big) \big(\sum_r b_{k,r} \big)^3 \Big)^{1/4} \le C \Big(\sum_{k=0}^{\infty} \sum_r a_r^4 b_{k,r} \Big)^{1/4} \le C \Big(\sum_r a_r^4 \Big)^{1/4}. \tag{18-4}$$

Finally for each r, an application of Hölder's inequality with exponents 8 and $\frac{8}{7}$ gives

$$a_r = \sum_{\beta:|E_\beta| \sim 2^r} 2^\beta |E_\beta|^{1/2} \max_{m,i} \Big(\frac{|E_\beta \cap \mathscr{C}_m^i|}{|E_\beta| + |\mathscr{C}_m^i|} \Big)^{s/4},$$

$$\le C 2^{r/2} \Big(\sum_{\beta:|E_\beta| \sim 2^r} 2^{4\beta} \max_{m,i} \Big(\frac{|E_\beta \cap \mathscr{C}_m^i|}{|E_\beta| + |\mathscr{C}_m^i|} \Big)^{2s} \Big)^{1/8} \Big(\sum_{\beta:|E_\beta| \sim 2^r} 2^{4\beta/7} \Big)^{7/8}$$

$$\le C \Big(\sum_{\beta:|E_\beta| \sim 2^r} 2^{4\beta} |E_\beta|^2 \max_{m,i} \Big(\frac{|E_\beta \cap \mathscr{C}_m^i|}{|E_\beta| + |\mathscr{C}_m^i|} \Big)^{s} \Big)^{1/8} \|f\|_2^{1/2},$$

since the sum of the finite series $\sum_{\beta:|E_\beta| \sim 2^r} 2^{4\beta/7}$ is comparable to its largest term.

Continuing now from (18-4), we have

$$\|f\|_{X_p}^8 \|f\|_2^{-4} \le C \sum_\alpha 2^{2\alpha} |E_\alpha| \cdot \sup_\alpha 2^{2\alpha} |E_\alpha| \max_{k,i} \Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \Big)^s$$

$$= C \|f\|_2^4 \cdot \sup_\alpha \Big(\big(2^{2\alpha} |E_\alpha| \|f\|_2^{-2}\big) \max_{k,i} \Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \Big)^s \Big)$$

$$\le C \|f\|_2^4 \cdot \sup_\alpha \Big(\big(2^{2\alpha} |E_\alpha| \|f\|_2^{-2}\big)^s \max_{k,i} \Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \Big)^s \Big)$$

for some $0 < s \le 1$.

It remains to show that

$$X := \sup_{\alpha}\Big(\big(2^{2\alpha}|E_\alpha|\|f\|_2^{-2}\big)\max_{k,i}\Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|}\Big)\Big) \le C \sup_{m,j}\Lambda_{m,j}(f)^r$$

for some positive exponent r. Choose an index α for which the supremum is attained up to a factor of at most 2. Then

$$\tfrac{1}{2}X \le \big(2^{2\alpha}|E_\alpha| \cdot \|f\|_2^{-2}\big)\max_{k,i}\Big(\frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|}\Big).$$

The right side is a product of two nonnegative factors, neither of which can exceed 1, so

$$2^{2\alpha}|E_\alpha|/\|f\|_2^2 \ge X/2 \quad \text{and there exist } k \text{ and } i \text{ such that} \frac{|E_\alpha \cap \mathscr{C}_k^i|}{|E_\alpha| + |\mathscr{C}_k^i|} \ge X/4.$$

Set $\mathscr{C} = \mathscr{C}_k^i$. We have $|E_\alpha| \ge 2^{-2\alpha-1}X\|f\|_2^2$, and since $|E_\alpha \cap \mathscr{C}| \le 2^{-\alpha}\int_{\mathscr{C}}|f|$,

$$|\mathscr{C}|^{-1}\int_{\mathscr{C}}|f| \ge 2^{\alpha}\frac{|E_\alpha \cap \mathscr{C}|}{|\mathscr{C}|} \ge 2^{\alpha}\frac{|E_\alpha \cap \mathscr{C}|}{|E_\alpha| + |\mathscr{C}|} \ge c2^{\alpha}X.$$

Also

$$\begin{aligned}|\mathscr{C}|^{-1}\int_{\mathscr{C}}|f| \ge 2^{\alpha}\frac{|E_\alpha \cap \mathscr{C}|}{|E_\alpha|}\cdot\frac{|E_\alpha|}{|\mathscr{C}|} \ge 2^{\alpha}\frac{|E_\alpha \cap \mathscr{C}|}{|E_\alpha| + |\mathscr{C}|}|\mathscr{C}|^{-1}|E_\alpha| &\ge c2^{\alpha}X|\mathscr{C}|^{-1}|E_\alpha| \\ &\ge c2^{\alpha}X|\mathscr{C}|^{-1}\cdot 2^{-2\alpha}\|f\|_2^2X = c2^{-\alpha}\|f\|_2^2X^2.\end{aligned}$$

Taking the geometric mean of these two bounds yields

$$\frac{|\mathscr{C}|^{-1}\int_{\mathscr{C}}|f|}{|\mathscr{C}|^{-1/2}\|f\|_2} \ge cX^{3/2},$$

which by the definitions of X and $\Lambda_{k,i}(f)$ is a bound of the desired form. □

Acknowledgement

We are indebted to Terence Tao for bringing the question to our attention, to Diogo Oliveira e Silva for useful comments on the exposition, and to an anonymous referee for exceptionally detailed and helpful comments.

References

[Andrews et al. 1999] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge University Press, 1999. MR 2000g:33001 Zbl 0920.33001

[Bégout and Vargas 2007] P. Bégout and A. Vargas, "Mass concentration phenomena for the L^2-critical nonlinear Schrödinger equation", *Trans. Amer. Math. Soc.* **359**:11 (2007), 5257–5282. MR 2008g:35190 Zbl 1171.35109

[Bennett et al. 2009] J. Bennett, N. Bez, A. Carbery, and D. Hundertmark, "Heat-flow monotonicity of Strichartz norms", *Anal. PDE* **2**:2 (2009), 147–158. MR 2010j:35418 Zbl 1190.35043

[Carneiro 2009] E. Carneiro, "A sharp inequality for the Strichartz norm", *Int. Math. Res. Not.* **2009**:16 (2009), 3127–3145. MR 2010h:35328 Zbl 1178.35090

[Christ 2011a] M. Christ, “On extremals for a Radon-like transform”, preprint, 2011. arXiv 1106.0728

[Christ 2011b] M. Christ, “Quasiextremals for a Radon-like transform”, preprint, 2011. arXiv 1106.0722

[Christ and Quilodrán 2010] M. Christ and R. Quilodrán, “Gaussians rarely extremize adjoint Fourier restriction inequalities for paraboloids”, preprint, 2010. To appear in *Trans. Amer. Math. Soc.* arXiv 1012.1346v1

[Christ and Shao 2012] M. Christ and S. Shao, “On the extremizers for an adjoint Fourier restriction inequality”, *Adv. Math.* **230** (2012), 957–97.

[Fanelli et al. 2011] L. Fanelli, L. Vega, and N. Visciglia, “On the existence of maximizers for a family of restriction theorems”, *Bull. Lond. Math. Soc.* **43**:4 (2011), 811–817. MR 2012g:42020 Zbl 1225.42012

[Foschi 2007] D. Foschi, “Maximizers for the Strichartz inequality”, *J. Eur. Math. Soc.* **9**:4 (2007), 739–774. MR 2008k:35389 Zbl 1231.35028

[Hundertmark and Zharnitsky 2006] D. Hundertmark and V. Zharnitsky, “On sharp Strichartz inequalities in low dimensions”, *Int. Math. Res. Not.* **2006** (2006), Art. ID 34080. MR 2007b:35277 Zbl 1131.35308

[Kunze 2003] M. Kunze, “On the existence of a maximizer for the Strichartz inequality”, *Comm. Math. Phys.* **243**:1 (2003), 137–162. MR 2004i:35006 Zbl 1060.35133

[Lions 1984a] P.-L. Lions, “The concentration-compactness principle in the calculus of variations, I: The locally compact case”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **1**:2 (1984), 109–145. MR 87e:49035a Zbl 0541.49009

[Lions 1984b] P.-L. Lions, “The concentration-compactness principle in the calculus of variations, II: The locally compact case”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **1**:4 (1984), 223–283. MR 87e:49035b Zbl 0704.49004

[Lions 1985a] P.-L. Lions, “The concentration-compactness principle in the calculus of variations, I: The limit case”, *Rev. Mat. Iberoamericana* **1**:1 (1985), 145–201. MR 87c:49007 Zbl 0704.49005

[Lions 1985b] P.-L. Lions, “The concentration-compactness principle in the calculus of variations, II: The limit case”, *Rev. Mat. Iberoamericana* **1**:2 (1985), 45–121. MR 87j:49012 Zbl 0704.49006

[Moyua et al. 1999] A. Moyua, A. Vargas, and L. Vega, “Restriction theorems and maximal operators related to oscillatory integrals in $\mathbb{R}^3$”, *Duke Math. J.* **96**:3 (1999), 547–574. MR 2000b:42017 Zbl 0946.42011

[Müller 1998] C. Müller, *Analysis of spherical symmetries in Euclidean spaces*, Applied Mathematical Sciences **129**, Springer, New York, 1998. MR 2001f:33004 Zbl 0884.33001

[Shao 2009] S. Shao, “Maximizers for the Strichartz and the Sobolev–Strichartz inequalities for the Schrödinger equation”, *Electron. J. Differential Equations* **13**:3 (2009), 1072–6691. MR 2010a:35032 Zbl 1173.35692

[Sogge 1993] C. D. Sogge, *Fourier integrals in classical analysis*, Cambridge Tracts in Mathematics **105**, Cambridge University Press, 1993. MR 94c:35178 Zbl 0783.35001

[Stein and Weiss 1971] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series **32**, Princeton University Press, 1971. MR 46 #4102 Zbl 0232.42007

[Xu 2000] Y. Xu, “Funk–Hecke formula for orthogonal polynomials on spheres and on balls”, *Bull. London Math. Soc.* **32**:4 (2000), 447–457. MR 2001g:33024 Zbl 1032.33005

Received 21 Jun 2010. Revised 11 Nov 2010. Accepted 22 Dec 2010.

MICHAEL CHRIST: mchrist@math.berkeley.edu
University of California, Berkeley, Department of Mathematics, Berkeley, CA 94720-3840, United States

SHUANGLIN SHAO: slshao@ima.umn.edu
Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455, United States

and

School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, United States

mathematical sciences publishers msp

Analysis & PDE

msp.berkeley.edu/apde

See inside back cover or msp.berkeley.edu/apde for submission instructions.

The subscription price for 2012 is US $140/year for the electronic version, and $240/year for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Analysis & PDE, at Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW™ from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
http://msp.org/
A NON-PROFIT CORPORATION
Typeset in LaTeX

ANALYSIS & PDE

Volume 5 No. 2 2012

2157-5045(2012)5:2;1-I