
Peer reviewed version

Link to published version (if available): 10.1093/imrn/rny117

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Oxford University Press at https://academic.oup.com/imrn/advance-article/doi/10.1093/imrn/rny117/5017349. Please refer to any applicable terms of use of the publisher.

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
$L^p$ NORMS OF EIGENFUNCTIONS ON REGULAR GRAPHS AND ON THE SPHERE

SHIMON BROOKS AND ETIENNE LE MASSON

Abstract. We prove upper bounds on the $L^p$ norms of eigenfunctions of the discrete Laplacian on regular graphs. We then apply these ideas to study the $L^p$ norms of joint eigenfunctions of the Laplacian and an averaging operator over a finite collection of algebraic rotations of the 2-sphere. Under mild conditions, such joint eigenfunctions are shown to satisfy for large $p$ the same bounds as those known for Laplace eigenfunctions on a surface of non-positive curvature.

1. Introduction

Let $(M, g)$ be an $n$-dimensional compact Riemannian manifold and $\Delta$ the Laplace-Beltrami operator. If $u$ is an eigenfunction, $\Delta u = \lambda^2 u$, Sogge proved [16] that

$$\|u\|_p \lesssim \lambda^{\sigma(n,p)} \|u\|_2,$$

with

$$\sigma(n,p) = \begin{cases} n \left( \frac{1}{2} - \frac{1}{p} \right) - \frac{1}{2} & \text{if } \frac{2(n+1)}{n-1} \leq p \leq \infty \\ \frac{n-1}{2} \left( \frac{1}{2} - \frac{1}{p} \right) & \text{if } 2 \leq p \leq \frac{2(n+1)}{n-1}. \end{cases}$$

(1.1)

These bounds are saturated on the sphere, in the case of high $p$ by the zonal spherical harmonics (concentration around points) and for low $p$ by Gaussian beams (concentration on closed geodesics).

Since the pioneering work of Sogge, the study of the $L^p$ norms of eigenfunctions has attracted a lot of interest as a way to understand how the geometry and the dynamics of the underlying space influences the shape of the eigenfunctions. In non-positive curvature, Berard showed [3] that the sup norm bound can be improved by a logarithmic factor, and more recently Hassel and Tacy [11] showed that this improvement can be maintained for all $L^p$ norms with $p > \frac{2(n+1)}{n-1}$, yielding

$$\|u\|_p \lesssim \frac{\lambda^{\sigma(n,p)}}{\sqrt{\log \lambda}} \|u\|_2.$$  

(1.2)

The field has expanded in many directions; see also eg. [18], [13], [8], and [12] for related results and their generalizations.

In this article, instead of Riemannian manifolds, we first consider the setting of regular graphs. These graphs can be seen as discrete analogues of hyperbolic surfaces and present similar characteristics to negatively curved manifolds. Our main result in this context is a bound on eigenfunctions of the discrete Laplacian (or adjacency matrix) analogous to (1.2). We then apply our graph methods to eigenfunctions satisfying some symmetry conditions on the sphere. More precisely, we look at joint eigenfunctions of the Laplacian and an averaging operator over a finite set of algebraic rotations satisfying mild non-degeneracy conditions, and show that for large $p$ such eigenfunctions satisfy the same bounds (1.2) as in the non-positive curvature case.

S.B. was supported by Israel Science Foundation grant 1119/13. E.L.M. was supported by the Marie Skłodowska-Curie Individual Fellowship grant 703162.
1.1. Eigenfunctions on regular graphs. Let $\mathcal{G}$ be a finite $q + 1$-regular graph, and $T_{q+1}$ its universal cover, the $q + 1$ regular tree. We will denote by $|\mathcal{G}|$ the number of vertices of the graph. We are interested in the eigenfunctions of the operator $T_q$ defined on $L^2(\mathcal{G})$ by

$$T_q f(x) = \frac{1}{\sqrt{q}} \sum_{d(x,y) = 1} f(y).$$

This operator is connected to the adjacency matrix $A$ and the Laplacian $\Delta$ of the graph by the formulas

$$\Delta = \frac{1}{q + 1} A - \text{Id} = \frac{\sqrt{q}}{q + 1} T_q - \text{Id},$$

so the eigenfunctions of $T_q$ coincide with the eigenfunctions of $A$ and $\Delta$.

In order to state the condition we require on the graph $\mathcal{G}$, we define, for any function $f : T_{q+1} \to \mathbb{C}$, the operator

$$\tilde{S}_n(f)(x) = \frac{1}{q^{n/2}} \sum_{d(x,y) = n} f(y).$$

Up to normalization, this is the averaging operator over spheres of radius $n$. Its projection to the graph will be denoted by $S_n$. We assume that for any $\delta > 0$, we have

$$\|S_n\|_{L^1 \to L^\infty} \lesssim q^{-1/2+\delta} n, \quad \text{for all } n \leq N. \quad (1.3)$$

where $N$ will be taken as large as possible. A natural assumption for us is that $N > c \log |\mathcal{G}|$ for some $c$ as $|\mathcal{G}| \to \infty$, as in this case Theorem 1.2 gives a bound analogous to (1.2). Here the size of the graph plays the role of the semiclassical parameter, as in [5].

Remark 1.1. Note that if we write

$$\delta_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise}, \end{cases}$$

then $\|S_n \delta_x\|_{L^\infty}$ is controlled by the number of cycles of length $2n$ through $x$. Condition (1.3) is always satisfied if $N$ is the radius of injectivity $\text{InjRad}(\mathcal{G})$, or equivalently if $2N$ is equal to the girth, and in this case we can take $\delta = 0$. As remarked in [6], if $\mathcal{G}$ is taken at random among the $q + 1$-regular graphs with fixed number of vertices, then the condition is satisfied for $N = c \log |\mathcal{G}|$, for some constant $c > 0$, with a probability tending to 1 as $|\mathcal{G}| \to \infty$. The main idea behind (1.3) is that we need to limit interferences when we propagate at distance $n \leq N$. In particular (1.3) will be violated if we propagate beyond logarithmic time; note for example that for the constant function $1$, condition (1.3) implies

$$q^n \lesssim q^{n/2} \|S_n 1\|_{L^\infty} \leq q^{\delta n} \|1\|_{L^1} = q^{\delta n} |\mathcal{G}|$$

for $\delta > 0$, so that a fortiori $N \leq \log_q |\mathcal{G}|$.

Theorem 1.2. Let $q \geq 2$ and $\mathcal{G}$ be a $q + 1$ regular graph satisfying (1.3) for some $N \geq \text{InjRad}(\mathcal{G})$, then for any eigenfunction $\psi$ of $T_q$ we have

$$\|\psi\|_p \lesssim_p \frac{\|\psi\|_2}{\sqrt{N}},$$

for any $p > 2$. 

Remark 1.3. If $G$ is a random regular graph of fixed degree and number of vertices, the theorem gives that for any $p > 2$,
\[ \|\psi\|_p \lesssim \|\psi\|_2 \sqrt{\frac{\log |G|}{|G|}} \]
almost surely when $|G| \to \infty$. However in the case of random regular graphs with a large fixed degree, much stronger results are shown in [2], namely
\[ \|\psi\|_\infty \lesssim \left(\frac{\log |G|}{|G|}\right)^{\delta} \|\psi\|_2 , \]
almost surely when $|G| \to \infty$, for some $\delta > 0$. On the other hand, our result holds for all graphs satisfying Condition (1.3), whereas the stronger bounds obtained from probabilistic methods can only be shown to hold almost surely.

There has been a growing interest recently in the study of eigenfunctions of the Laplacian on regular graphs seen both as independent combinatorial objects (deterministic [6, 1] and random [2]) and as graphs arising from symmetries on manifolds [7]. A first result of delocalization of eigenfunctions was given in [6]. Although it is not explicitly derived in the paper, the proof gives a bound on the sup norm of eigenfunctions, but the uniform bound obtained directly from there is only
\[ \|\psi\|_\infty \lesssim \frac{\|\psi\|_2}{N^{1/4}} , \]
Our methods are similar to [6] — based on harmonic analysis on regular trees — and combine an adaptation of the argument of [6] to some ideas of [11]. They allow us to prove a better bound and generalize it to all $L^p$ norms for any $p > 2$.

1.2. Eigenfunctions on the sphere. Eigenfunctions of the Laplacian on the round 2-sphere lie at the opposite extreme from hyperbolic surfaces: here the dynamics are stable, the spectrum has large multiplicities, the property of quantum ergodicity fails [9], and the Sogge $L^p$ bounds (1.1) are achieved [15, 17]. On the other hand, the large multiplicities allow for different bases that exhibit different properties, and in joint work with E. Lindenstrauss [5] we showed that quantum ergodicity does hold for bases of joint eigenfunctions of the Laplacian and an averaging operator over a finite set of rotations (stronger results for random bases were already proved in [21] and [20]). This followed from a quantum ergodicity result on large regular graphs (first proved in [1]), by essentially using the graph structure induced from the free subgroup of rotations, to transfer techniques from the graph setting to the sphere $S^2$. In this sense, our result on the sphere is in the same spirit as [5], as we take advantage of the additional symmetries to transfer our $L^p$ methods on graphs to get the Hassel-Tacy bounds (1.2) on the sphere. We remark that in order to get a logarithmic improvement, we will have to make some assumptions on the set of rotations, most notably that they not get too close to each other (in a very quantitative way); this condition is satisfied when the rotations are algebraic (see Lemma 4.2).

So for $M \geq 2$, let $g_1, \ldots, g_M$ be a finite set of rotations in $SO(3)$ with algebraic entries, that generate a free subgroup. We further assume that the stabilizer of any $x \in S^2$ is either trivial or cyclic — i.e., its intersection with words of length $n$ does not grow exponentially in $n$ — in order for the images $\{g.x : g$ is a word of length $n\}$ to conform with Condition (1.3). We define the operator $T_q$ acting on $L^2(S^2)$ by
\[ T_q f(x) = \frac{1}{\sqrt{q}} \sum_{j=1}^{M} \left( f(g_j x) + f(g_j^{-1} x) \right) , \]
with \( q = 2M - 1 \).

**Theorem 1.4.** Let \( \psi_s \) be an \( L^2 \)-normalized eigenfunction of the positive Laplacian \( \Delta \) on the sphere \( S^2 \)

\[
\Delta \psi_s = s(s+1)\psi_s
\]

Assume that \( \psi_s \) is also an eigenfunction of an averaging operator \( T_q \) for a finitely-generated free algebraic subgroup of \( SO(3) \) such that every non-trivial stabilizer of \( x \in S^2 \) is cyclic. Then we have for \( p > 6 \)

\[
\| \psi_s \|_p \lesssim \frac{s^{\frac{1}{2}} - \frac{2}{p}}{\sqrt{\log s}}.
\]

Note that this matches the best-known bounds for surfaces of non-positive curvature (1.2), proved by Hassell and Tacy [11].

**Remark 1.5.** In the setting of congruence surfaces, spectral multiplicities are expected to be uniformly bounded, and thus it is natural to focus attention to the basis of Hecke-Maass forms. But on the sphere, where spectral multiplicities are large, the hypotheses of Theorem 1.4 are much more restrictive, and should not be regarded as generic or representative; rather, they are indicative of the type of dynamical structure necessary in order to apply the current technology in a deterministic setting.

**Remark 1.6.** In order to keep the exposition concise and self-contained, we will give a simplified argument that proves Theorem 1.4 for \( p > 8 \). The more general result for \( p > 6 \) is identical, but requires Sogge’s more refined argument from [15, 17] in place of the “baby” version we include that appeals directly to Young’s inequality; we outline these refinements in Section 4.3.

The paper is organized as follows. In Section 2, we prove Theorem 1.2 in the case of Ramanujan graphs. These are graphs satisfying an additional condition on the spectrum of the adjacency matrix (a maximal spectral gap), which simplifies the proof. This allows us to show the main ideas of the argument, in a case with minimal technical difficulties. Moreover, the result is stronger for these graphs, as we only require a weakened version of Condition (1.3), and is therefore interesting in its own right. The general proof of Theorem 1.2, without the Ramanujan assumption, is addressed in Section 3. We then prove Theorem 1.4 in Section 4 by combining the ideas of the previous sections with estimates on spectral projection operators.

**Acknowledgements.** We thank Melissa Tacy for helpful discussions, and in particular for patiently walking us through some of the relevant background for Section 4.

2. The case of Ramanujan graphs

We parametrize the spectrum of \( T_q \) by \( \lambda = 2 \cos \theta_\lambda \), with

\[
\lambda \in \left[ -2 \cosh \left( \frac{\log q}{2} \right), 2 \cosh \left( \frac{\log q}{2} \right) \right].
\]

It can be divided into two parts: the **tempered spectrum** is the part contained in the interval \([-2, 2]\), and the **untempered spectrum** is the part lying outside this interval (corresponding to complex values of \( \theta_\lambda \)).

In order to outline the main argument of the proof, we first consider the case of Ramanujan graphs. These are graphs with an empty untempered spectrum (except for the eigenvalue \( 2 \cosh \left( \frac{\log q}{2} \right) \) corresponding to the constant eigenfunctions). In
this case we can weaken the girth assumption (1.3) and only ask that there exist \( \beta > 0 \) such that

\[
\|S_n\|_{L^1 \rightarrow L^\infty} \lesssim q^{-\beta n}, \quad \text{for all } n \leq N. \tag{2.1}
\]
as in [6]. Note that although the famous construction of Ramanujan graphs of [14] gives graphs with large girth, automatically satisfying Condition (2.1) for \( N \gtrsim \log |G| \), this is not a general property (see [10]) and we therefore need the condition.

We will consider the operator \( S_n \Pi \), where \( \Pi \) is the orthogonal projection on the tempered spectrum. In the case of Ramanujan graphs we have \( \Pi f(x) = f(x) - \bar{f} \), where \( \bar{f} = \frac{1}{|\mathcal{G}|} \sum_{y \in \mathcal{G}} f(y) \). In the proof of the general theorem, we will need to subtract terms involving other untempered eigenfunctions and control their norm, which requires the stronger assumption (1.3).

Note that Theorem 1.2 holds trivially for the constant function, since

\[
\|1\|_p = \frac{1}{|\mathcal{G}|} |\mathcal{G}|^{1/p} = |\mathcal{G}|^{1/p - 1/2}\]

(In fact, one easily disposes with all untempered eigenfunctions, see Lemma 3.1.) Thus we may assume that \( \psi \) is orthogonal to constants; i.e., that \( \Pi \psi = \psi \).

**Lemma 2.1.** Assuming (2.1), there exists \( \beta_p > 0 \) such that we have the estimate

\[
\|S_n \Pi\|_{L^{p'} \rightarrow L^p} \lesssim q^{-\beta_p n}, \quad \text{for all } n \leq N,
\]

for any conjugate exponents \( 1 \leq p' < 2 < p \leq \infty \), i.e. such that \( \frac{1}{p} + \frac{1}{p'} = 1 \).

**Proof.** Note that for any \( f \) and \( n \leq N \),

\[
\|S_n \Pi f\|_\infty = \|S_n f - S_n \bar{f}\|_\infty \\
\leq \|S_n f\|_\infty + \|S_n \bar{f}\|_\infty \\
\lesssim \|S_n\|_{L^1 \rightarrow L^\infty} (\|f\|_1 + \|\bar{f}\|_1) \\
\lesssim q^{-\beta n} \|f\|_1.
\]

We deduce that

\[
\|S_n \Pi\|_{L^1 \rightarrow L^\infty} \lesssim q^{-\beta n}, \quad \text{for all } n \leq N.
\]

On the other hand, it is well known that \( S_n \) can be expressed as a polynomial in \( T_q \), namely \( S_n = \frac{q + 1}{q} q^{n/2} F(T_q, n) \), where \( F(2 \cos \theta, n) = \phi_0(n) \) is the spherical function

\[
\phi_0(n) = q^{-n/2} \left( \frac{2}{q + 1} \cos(n \theta) + \frac{q - 1}{q + 1} \sin(n + 1) \theta \right) \tag{2.2}
\]

Equivalently, the eigenfunctions of \( T_q \) and \( S_n \) coincide and the spectrum of \( S_n \) is the set \( \{ \frac{q + 1}{q} q^{n/2} \phi_0(j), j = 1, \ldots, |\mathcal{G}| \} \), where \( \{ 2 \cos(\theta_j) \}_{j=1, \ldots, |\mathcal{G}|} \) is the spectrum of \( T_q \). The spectrum of \( S_n \Pi \) is obtained purely from parameters \( \theta \) in the tempered spectrum of \( T_q \), i.e. \( \theta \in \mathbb{R} \), and is therefore bounded by \( n \). We have

\[
\|S_n \Pi\|_{L^2 \rightarrow L^2} = O(n).
\]

By interpolation we obtain

\[
\|S_n \Pi\|_{L^{p'} \rightarrow L^p} \lesssim q^{-\beta_p n}, \quad \text{for all } n \leq N.
\]

for any conjugate exponents \( \frac{1}{p} + \frac{1}{p'} = 1 \) with \( p > 2 \), by setting eg. \( \beta_p = \frac{1}{2} \beta(1 - \frac{2}{p}) \).

We now let \( \lambda = 2 \cos \alpha \), with \( \alpha \in [0, \pi] \), be the eigenvalue of a non-constant eigenfunction \( \psi \) (in our previous notation \( \alpha = \theta_\lambda \)). The idea of the proof is to define an operator that localizes spectrally near the eigenvalue \( \lambda \) and to bound its \( L^2 \rightarrow L^p \) norm for \( p > 2 \). The bound will be obtained by a \( TT^* \) argument similar to what was used in [17, 11].
We define for $N$ even

$$W_{N, \alpha} = \sum_{n=1}^{N/2} \cos(2n\alpha) P_{2n}(T_q/2)$$

where $P_n(\cos \theta) = \cos(n\theta)$ is the Chebyshev polynomial of the first kind. We can see $P_n(T_q/2)$ as a discrete wave propagator as suggested in [6]. The operator $W_{N, \alpha}$ is then reminiscent of the spectral cluster operators usually considered in the study of $L^p$ norms of eigenfunctions (see in particular [11]).

The spectral localization, or at least the emphasis of the eigenvalue $\lambda$ is based on the following lemma from [5], whose proof we reproduce for the convenience of the reader.

**Lemma 2.2.** For every $\alpha \in \mathbb{R}$ and $N \geq 10$, we have

$$\sum_{n=1}^{N} \cos(n\alpha)^2 \geq 0.3N$$

**Proof.** As cosine is $2\pi$-periodic and even, it is sufficient to consider $\alpha \in [0, \pi]$. We write

$$S_T := \sum_{n=1}^{T} \cos(n\alpha)^2.$$

Then

$$\frac{1}{T} S_T = \frac{1}{T} \sum_{n=1}^{T} \left( \frac{1}{2} + \frac{1}{4} e^{2i\alpha n} + \frac{1}{4} e^{-2i\alpha n} \right)$$

$$= \frac{2T - 1}{4T} + \frac{1}{4T} \sum_{n=-T}^{T} e^{2i\alpha n}$$

$$= \frac{2T - 1}{4T} + \frac{\sin[(2T + 1)\alpha]}{4T \sin \alpha}.$$

Now

$$\min_{\alpha \in [0, \pi]} \frac{\sin[(2T + 1)\alpha]}{4T \sin \alpha} = \min_{\alpha \in [0, \pi/2]} \frac{\sin[(2T + 1)\alpha]}{4T \sin \alpha} \geq -\frac{1}{4T \sin(\pi/(2T + 1))},$$

since $|\sin \alpha|$ is monotone increasing in $[0, \pi/2]$ and $\sin(2T + 1)\alpha > 0$ for $0 < \alpha < \pi/(2T + 1)$.

Hence

$$\frac{1}{T} S_T \geq \frac{1}{4T} \left( 2T - 1 - \frac{1}{\sin(\pi/(2T + 1))} \right).$$

For $T \geq 10$ the right-hand side of the inequality is $\geq 0.3$. \qed

Using Lemma 2.2 and the fact that $\Pi \psi = \psi$, we get that

$$W_{N, \alpha} \psi = \sum_{n=1}^{N/2} \cos(2n\alpha) P_{2n}(\cos \alpha) \psi$$

$$= \sum_{n=1}^{N/2} \cos^2(2n\alpha) \psi = \sum_{n=1}^{N/2} \cos^2(n[2\alpha]) \psi$$

and therefore

$$\|W_{N, \alpha} \psi\|_p \geq 0.3N \|\psi\|_p.$$

Now we can write

$$N \|\psi\|_p \lesssim \|W_{N, \alpha} \psi\|_p \leq \|W_{N, \alpha}\|_{L^2 \to L^p} \|\psi\|_2.$$
Hence the proof of Theorem 1.2 is reduced to the following proposition.

**Proposition 2.3.** We have 
\[ \|W_{N,\alpha}\|_{L^2 \to L^p} = O(\sqrt{N}), \]
for any \( 2 < p \leq \infty \).

To prove Proposition 2.3 we first recall the following lemma as stated in [6].

**Lemma 2.4.** Fix a point \( o \in T_{q+1} \) and write \( |x| = d(o, x) \) for any \( x \in T_{q+1} \). Let \( \delta_o \) be the \( \delta \)-function supported at \( o \), and \( n \) a positive even integer. Then
\[ P_n(T_q/2)\delta_o(x) = \begin{cases} 0 & |x| \text{ odd or } |x| > n \\ \frac{1-q}{2q^{n/2}} & |x| < n \text{ and } |x| \text{ even} \\ \frac{1}{2q^{n/2}} & |x| = n \end{cases} \]

We then have the estimate (see Corollary 2 in [6])

**Lemma 2.5.** Let \( N \) be as in Condition (1.3). For any conjugate exponents \( 1 \leq p' < 2 < p \leq \infty \) there exists \( \beta > 0 \) such that
\[ \|P_n(T_q/2)\Pi\|_{L^{p'} \to L^p} \lesssim q^{-\beta n}, \]
for any \( n \leq N \).

**Proof.** The proof is identical to Corollary 2 in [6] and based on the fact that we can write
\[ P_n(T_q/2) = \sum_{k=0}^{n/2-1} \frac{1-q}{2q^{n/2}} q^k S_{2k} + \frac{1}{2} S_n. \]

We then use the estimate of Lemma 2.1 \( \square \)

We are now ready to prove the proposition.

**Proof of Proposition 2.3.** We use a \( TT^* \) argument, or in other words the fact that
\[ \|T\|_{L^2 \to L^p}^2 = \|T^*\|_{L^{p'} \to L^2}^2 = \|TT^*\|_{L^p \to L^p} \]
for conjugate exponents \( \frac{1}{p} + \frac{1}{p'} = 1. \)

First note that
\[ W_{N,\alpha}W_{N,\alpha}^* = \sum_{n=1}^N \sum_{k=1}^N \cos(n\alpha) \cos(k\alpha) P_n(T_q/2) P_k(T_q/2) \Pi \]

Now we use Lemma 2.5 in addition to the formula
\[ P_n P_k = \frac{1}{2} (P_{n+k} + P_{|n-k|}) \]
satisfied by the Chebyshev polynomials, to get
\[ \|W_{N,\alpha}W_{N,\alpha}^*\|_{L^{p'} \to L^p} \leq \sum_{n=1}^N \sum_{k=1}^N \|P_n(T_q/2) P_k(T_q/2) \Pi\|_{L^{p'} \to L^p} \]
\[ \leq \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^N \left( \|P_{n+k}(T_q/2) \Pi\|_{L^{p'} \to L^p} + \|P_{|n-k|}(T_q/2) \Pi\|_{L^{p'} \to L^p} \right) \]
\[ \leq \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^N q^{-\beta n} q^{-\beta k} + \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^N q^{-\beta |n-k|} \]
\[ \leq O(1) + \sum_{n=1}^N \sum_{k=n}^N q^{-\beta (k-n)} = O(N) \]
By (2.3), we deduce
\[ \|W_{N,\alpha}\|_{L^2 \to L^p} = O(\sqrt{N}). \]

Using Proposition 2.3, we finally obtain for any \(2 < p \leq \infty\)
\[ \|\psi\|_p \lesssim_p \frac{1}{\sqrt{N}} \|\psi\|_2. \]

3. Proof of Theorem 1.2

The proof for the general case is similar to the case of Ramanujan graphs. Here we need the stronger assumption (1.3). Denote by \(K_k(x,y)\) the kernel of the operator \(S_k\). We can then rewrite (1.3) as
\[ \sup_{x,y} |K_k(x,y)| \lesssim \delta \left( q^{-1/2} + \delta \right)^k \]
for all \(k \leq N\). We denote \(\Pi_\varepsilon\) to be the projection to the part of the spectrum in \([-2-\varepsilon, 2+\varepsilon]\), that is in an \(\varepsilon\)-neighborhood of the tempered spectrum. The goal is to bound \(\|S_k \Pi_\varepsilon\|_{L^1 \to L^\infty}\) to prove the analogue of Lemma 2.1; Theorem 1.2 holds trivially for untempered \(\psi\) as we will see below in Lemma 3.1, like the case of the constant function in section 2 above. In particular, we may safely assume for every \(\varepsilon > 0\) that \(\Pi_\varepsilon \psi = \psi\).

More precisely, assumption (1.3) means that the sup-norm of an untempered eigenfunction decays exponentially with \(N\). In fact — and this is the key point — any function in the span of untempered eigenfunctions will also have a sup-norm that decays exponentially:

**Lemma 3.1.** Let
\[ f(x) = \sum_j \phi_j(x) \]
where the \(\{\phi_j\}\) are mutually orthogonal untempered eigenfunctions, each of eigenvalue \(> 2 + \varepsilon\) in absolute value; i.e., \(\Pi_\varepsilon f = 0\). Then there exists \(\delta(\varepsilon) > 0\) such that
\[ \|f\|_\infty \lesssim \varepsilon q^{-\delta(\varepsilon)k} \|f\|_2 \]
for every \(k \leq N\) satisfying (1.3).

**Proof.** Let \(x\) be a point at which \(|f(x)| = \|f\|_\infty\). We may assume without loss of generality that \(f(x) > 0\), and that each \(\phi_j(x) > 0\) — since otherwise we could omit the negative ones and increase \(\|f\|_\infty\) while reducing \(\|f\|_2\).

Recall from (2.2) that if \(\phi\) is an eigenfunction of \(\Delta\) of eigenvalue \(\lambda = 2 \cos \theta\), then
\[ S_k \phi = \frac{q+1}{q} q^{n/2} F(2 \cos \theta, k) \phi, \]
where
\[ F(2 \cos \theta, k) = q^{-n/2} \left( \frac{2}{q+1} \cos(k \theta) + \frac{q-1}{q+1} \frac{\sin(k+1) \theta}{\sin \theta} \right) \]
and \(F(2 \cos \theta, k)\) is the spherical function. For simplicity we first consider \(k\) even. In this case all eigenvalues of \(S_k\) are positive, which can be seen from the relation \(F(-2 \cos \theta, k) = (-1)^k F(2 \cos \theta, k)\).

We write
\[
S_k f(x) = \sum_j \lambda_{j,k} \phi_j(x) \geq \lambda_{\varepsilon,k} \sum_j \phi_j(x)
\]
\[ \geq \lambda_{\varepsilon,k} |f(x)| \]
where $\lambda_{\epsilon,k}$ is the eigenvalue for $S_k$ acting on an untempered eigenfunction of $T_q$-eigenvalue $2 + \epsilon$, and the $\lambda_{j,k} > \lambda_{\epsilon,k}$ denote the respective eigenvalues of $S_k$ on each untempered $\phi_j$. We have

$$
\lambda_{\epsilon,k} = \frac{2}{q} \cos(k\theta) + \frac{q - 1}{q} \sin(k + 1) \frac{\sin \theta}{\sin \theta}
$$

where $\theta$ is defined by the relation $2 \cos(\theta) = 2 + \epsilon$, i.e. $\theta = -i \arccosh(1 + \epsilon/2)$. We see that $\lambda_{\epsilon,k}$ grows exponentially in $k$, with exponent growth determined by $\epsilon$; i.e.

$$
\log_q \lambda_{\epsilon,k} = C(\epsilon) \cdot k + O(1)
$$

On the other hand, we have that for any $\delta > 0$,

$$
||K_k(x, \cdot)||_2 \lesssim \delta q^{k\delta/2}
$$

by assumption (1.3), and so we see from Cauchy-Schwarz inequality

$$
|S_k f(x)| \leq ||K_k(x, \cdot)||_2 ||f||_2 \lesssim \delta q^{k\delta/2} ||f||_2.
$$

Combining this with (3.1) gives

$$
\lambda_{\epsilon,k} |f(x)| \leq q^{k\delta/2} ||f||_2.
$$

Choosing $\delta(\epsilon) < C(\epsilon)$ small enough that $k\delta(\epsilon) < \log_q \lambda_{\epsilon,k}$, we get

$$
|f(x)| \lesssim_{\delta(\epsilon)} q^{-k\delta(\epsilon)/2}
$$

as required. Further adjusting $\delta(\epsilon)$ if necessary, we can guarantee that the inequality extends to odd values of $k$ as well. □

The above Lemma will allow us to handle situations where there are many untempered eigenfunctions, because any linear combination of them still has sup-norm that decays exponentially relative to the $L^2$-norm. Now we need to exploit the $L^2$-norm in order to bound $S_n \Pi_\epsilon$ from $L^1 \to L^\infty$. We note that for any function on the graph, we have

$$
||f||_2 \leq ||f||_1,
$$

Indeed, without loss of generality we may assume that $||f||_1 = 1$. This means in particular that $|f(x)| \leq 1$ for all $x$ in the graph, and so $|f(x)|^2 \leq |f(x)|$ for every $x$, yielding

$$
||f||_2^2 = \sum_x |f(x)|^2 \leq \sum_x |f(x)| \leq ||f||_1 = 1
$$

Taking square-roots gives $||f||_2 \leq 1$ as required. This fact will be combined with Lemma 3.1 to prove the following:

**Lemma 3.2.** Under the assumption (1.3), we have that for every $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$
||S_n \Pi_\epsilon||_{L^1 \to L^\infty} \lesssim_{\epsilon} q^{-n/4}
$$

for $n \leq \delta(\epsilon) \cdot N$, where $N$ is from assumption (1.3).

**Proof.** Consider $||f||_1 = 1$, and write $f = \Pi_\epsilon f + f_{\text{untemp}}$ as the decomposition of $f$ into an “$\epsilon$-tempered” piece and a second piece consisting of all sufficiently untempered spectral components of $T_q$-eigenvalue greater than $2 + \epsilon$ in absolute-value.
We have
\[ \|S_n \Pi f\|_\infty = \|S_n f - S_n f_{\text{untemp}}\|_\infty \leq \|S_n f\|_\infty + \|S_n f_{\text{untemp}}\|_\infty \leq q^{-n/4} \|f\|_1 + \|S_n f_{\text{untemp}}\|_\infty \lesssim q^{-n/4} + \|S_n f_{\text{untemp}}\|_\infty \]
by the assumption (1.3) on the kernel which guarantees that
\[ \|S_n\|_{L^1 \to L^\infty} \lesssim q^{-n/4} \|f\|_2 \]
So it remains to show that
\[ \|S_n f_{\text{untemp}}\|_\infty \lesssim q^{-n/4} \]
But since \( \|f\|_1 = 1 \), we have by (3.2) that \( \|f\|_2 \leq 1 \), which means in particular that \( \|f_{\text{untemp}}\|_2 \leq 1 \). And since \( S_n \) preserves the eigenspaces, \( S_n f_{\text{untemp}} \) also belongs to the kernel of \( \Pi_\varepsilon \), and thus we can apply Lemma 3.1 to show that
\[ \|S_n f_{\text{untemp}}\|_\infty \lesssim \varepsilon q^{-n/4} \lambda_{\varepsilon(p),n} \leq q^{n/2} \]
where we have used the fact that \( \|S_n\|_{L^2 \to L^2} \) is bounded\(^1\) by \( q^{n/2} \) and the claim follows as long as \( n \leq \delta(\varepsilon)N \).

\[ \square \]

The rest of the argument is as follows. Fix \( p > 2 \), and \( \varepsilon(p) \) to be chosen later. We assume \( N \gtrsim \log |\mathcal{G}| \) and set \( n = \delta(\varepsilon(p))N \). The factor \( \delta(\varepsilon(p)) \) will appear in the final estimate for \( \|\psi\|_p/\|\psi\|_2 \).

We have
\[ \|S_n \Pi_{\varepsilon(p)}\Pi f\|_\infty \lesssim \varepsilon(p) q^{-n/4} \|S_n f_{\text{untemp}}\|_2 \lesssim \delta(\varepsilon(p)) \lambda_{\varepsilon(p),n} \leq q^{n\varepsilon(p)} \]
We pick \( \varepsilon(p) \) small enough that interpolation gives an exponential decay bound
\[ \|S_n \Pi_{\varepsilon(p)}\Pi f\|_{L^p' \to L^p} \lesssim \varepsilon(p) q^{-0.01n} \]
for all \( n < \delta(\varepsilon(p))N \), and thus in the \( TT^* \) argument we will get a bound of
\[ \frac{\|\psi\|_p}{\|\psi\|_2} \lesssim \varepsilon(p) \frac{1}{\sqrt{\delta(\varepsilon(p))N}} \lesssim p \frac{1}{\sqrt{N}} \]
for all eigenfunctions \( \psi \), with implied constant determined by \( p \).

4. Joint Eigenfunctions on the Sphere

In this section we apply the above arguments to the case of joint eigenfunctions of the Laplacian and an averaging operator over a free group of algebraic rotations. For simplicity, we will give a self-contained proof of the result for \( p > 8 \); to extend to the full strength of Theorem 1.4, one couples the argument below with Sogge’s “freezing” method to exploit stationary phase cancellations and reduce the estimate to a one-dimensional Young’s inequality, instead of the wasteful two-dimensional Young’s inequality we apply below (see Section 4.3).

\(^1\)Here, a uniform spectral gap could improve the bounds a bit.
We let $\mathcal{H}_s$ be the eigenspace of spherical harmonics of eigenvalue $s(s + 1)$, $s \in \mathbb{N}$, and $Z_s \in \mathcal{H}_s$ the zonal spherical harmonic centered at the North pole 0. These have the reproducing property that
\[
Z_s \ast \psi = \psi \tag{4.1}
\]
for any $\psi \in \mathcal{H}_s$, where the convolution is defined by
\[
(Z_s \ast \psi)(x) = \int_{S^2} Z_s(d(x, y))\psi(y)\,d\sigma(y),
\]
and $\sigma$ denotes the area measure on the sphere. We have $Z_s(x) = Z_s^0(d(0, x))$ for a function $Z_s^0$ on $[0, \pi]$. We will abuse notation and drop the extra decoration, and see $Z_s$ itself as a function on $[0, \pi]$. It can be written in terms of the Legendre polynomial $L_s$ of degree $s$ as
\[
Z_s(t) = \frac{2s + 1}{4\pi} L_s(\cos t),
\]
for $t \in [0, \pi]$. We will use standard bounds on Legendre polynomials [19, Theorem 7.3.3], which yield the estimate
\[
Z_s(t) \lesssim \min\{s, \sqrt{s/\sin t}\}. \tag{4.2}
\]

4.1. $L^p$ bounds for spherical harmonics, $p > 8$. In this section we motivate our argument through a simplified proof of Sogge’s bounds ([15]) for spherical harmonics, in the easier case $p > 8$ where we may simply apply Young’s inequality.

The reproducing property (4.1), together with orthogonality of eigenspaces, means that convolution with $Z_s$ is the projection to $\mathcal{H}_s$. Thus it is the norm of this convolution $\chi_s$: $f \mapsto Z_s \ast f$ as an operator from $L^2(S^2) \to L^p(S^2)$ that we wish to estimate. Since this convolution is a projection, we could equally well estimate the norm from $L^{p'}(S^2) \to L^p(S^2)$ for $\frac{1}{p} + \frac{1}{p'} = 1$ and take the square root, since
\[
\|\chi_s\|_{L^2 \to L^p}^2 = \|\chi_s \chi_s^*\|_{L^{p'} \to L^p} = \|\chi_s\|_{L^{p'} \to L^p}
\]

Lemma 4.1 ($L^p$ bounds for spherical harmonics, $p > 8$). We have
\[
\|\chi_s\|_{L^2 \to L^p} \lesssim s^{\frac{1}{2} - \frac{2}{p}}.
\]

Proof. Since we have explicit bounds on the kernel $Z_s$ of $\chi_s$, let us apply Young’s inequality directly to estimate $\|\chi_s\|_{L^{p'} \to L^p}$:
\[
\|\chi_s \psi\|_{L^p} \leq \|Z_s\|_{L^{p/2}} \cdot \|\psi\|_{L^{p'}}.
\]
The bound (4.2) gives
\[
\|Z_s\|_{L^{p/2}} \lesssim \left( \int_0^{s^{-1}} |Z_s(t)|^{p/2} \sin t \, dt \right)^{2/p} + \left( \int_{s^{-1}}^{\pi} |Z_s(t)|^{p/2} \sin t \, dt \right)^{2/p} \\
+ \left( \int_{s^{-1}}^{\pi-s^{-1}} |Z_s(t)|^{p/2} \sin t \, dt \right)^{2/p}
\]
\[
\lesssim \left( \frac{s^{p/2}}{s-2} \right)^{2/p} + \left( \frac{s^{p/4}}{t^{p/4+1}} \right)^{2/p}
\]
\[
\lesssim s^{1-\frac{4}{p}} + s^{1/2} \left( t^{-\frac{2}{p}+\frac{1}{s-1}} \right)^{2/p}
\]
\[
\lesssim s^{1-\frac{4}{p}} + s^{1/2} \left( \frac{p}{s+2} \right)^{2/p}
\]
\[
\lesssim s^{1-\frac{4}{p}} + s^{1/2} s^{-\frac{1}{2}+\frac{4}{p}}
\]
\[
\lesssim s^{1-\frac{4}{p}}
\]
\[(4.3)\]
as long as \(p > 8\). Substituting this back into Young’s inequality above gives
\[
\|\chi_s \psi\|_{L^p} \lesssim s^{1-\frac{4}{p}} \|\psi\|_{L^{p'}}
\]
whereby
\[
\|\chi_s\|^2_{L^2 \to L^p} = \|\chi_s\|_{L^{p'} \to L^p} \lesssim s^{1-\frac{4}{p}}
\]
as required. \(\square\)

4.2. The Improvement for Joint Eigenfunctions. In order to apply the results of the previous sections, we wish to average together many copies of the kernel centered at different points determined by the averaging operator. To this end, note that the essential contribution to the norm of \(\chi_s\) comes from the \(s^{-1+\delta}\) neighborhood of the center point; see precise statement in (4.5) below. Thus the contribution of points outside the radius of \(s^{-1+\delta}\) is negligible, and we may safely combine copies of the kernel \(Z_s\), centered at points that are \(s^{-1+\delta}\)-separated, without affecting the estimates.

Thus we must show that for our group of algebraic rotations, the image of any point under a word of length \(n\) in the generators consists of points separated by a distance of \(s^{-1+\delta}\). This is where the algebraicity will come in, as in [5]. However, we cannot expect this to hold for all images at all points, since every rotation fixes a pair of antipodal points on the sphere, and so e.g. any fixed point of one of the generators will be fixed by powers of the generator, and thus will be fixed by (at least) \(n\) words of length \(\leq n\). This is where we will need to use the assumption on the stabilizers, to ensure that we have only linear growth of the non-trivial stabilizers, and do not have exponential growth (which would violate Condition (1.3)).

First, we need the following Lemma, essentially copied from [5, Lemma 4.2]:

**Lemma 4.2.** There exists \(c > 0\) depending only on the generating set, such that for all \(N < c \log s\) and all \(x \in S^2\), there exist at most 2 words \(g\) of length \(N\) in our generators such that \(d(x, g.x) \leq 2s^{-1/4}\).

As we mentioned above, we have to allow some “bad” words, since if \(x\) is stabilized by some word \(g\), then powers of \(g\) will also stabilize \(x\). Moreover a small neighborhood (e.g., a ball of radius \(s^{-1}\)) of the fixed point will also stay close to itself under these
powers of $g$. But since the stabilizer of any such $x$ is cyclic, there can be only 2 words of a given length, $g^n$ and $g^{-n}$, that fix $x$ and have the same axis of rotation.

**Proof of Lemma 4.2.** Let $K$ be a finite extension of $\mathbb{Q}$ containing all entries of the matrices $g_1, \ldots, g_M$. Recall the notion of (logarithmic) height of an algebraic number $\alpha \in K$ from e.g. [4, Ch. 1]. What is important for us is that it is a nonnegative real number measuring the complexity of nonzero $\alpha \in K$ (e.g. for a rational number given in lowest terms by $p/q$ we have that $h(p/q) = \log \max(p, q)$) with the following properties:

- $h(\alpha\beta) \leq h(\alpha) + h(\beta)$
- $h(\alpha + \beta) \leq h(\alpha) + h(\beta) + \log 2$ (cf. [4, §1.5.16])
- for any embedding $\iota : K \rightarrow \mathbb{C}$ and any algebraic number $\alpha \neq 0$ we have that $|\iota(\alpha)| \geq e^{-|K:Q|h(\alpha)}$ (cf. [4, §1.5.19]).
- $h(\alpha^{-1}) = h(\alpha)$

It follows that if we set, for $g \in \text{SO}(3, \mathbb{Q})$, the height $h(g)$ of $g$ to be the maximum of the heights of its coordinates, then for any $g_1, g_2 \in \text{SO}(3, \mathbb{Q})$ we have

$$h(g_1 g_2) \leq h(g_1) + h(g_2) + O(1)$$
$$h(g_1^{-1}) \leq O(h(g_1) + 1).$$

Thus if $w(g_1, \ldots, g_M)$ is any word of length $N$ in the given generating set

$$h(w(g_1, \ldots, g_M)) \lesssim N(1 + \max\{h(g_1), \ldots, h(g_M)\}).$$

Since $g_1, \ldots, g_M$ generate a free group, it follows from this and the above basic property of heights, that for any reduced word $w$ of length $N$ if $g = w(g_1, \ldots, g_M)$ then $\|g - 1\| \geq C^{-N}$ for some $C$ depending only on the generating set, in other words $g$ is a rotation of angle $\theta$ around its axis with $|\theta| \geq C^{-N}$. Moreover, since the commutator of two words of length $N$ is a word of length at most $4N$, the commutators satisfy a similar bound (up to adjusting the constant), and therefore we deduce that the distinct axes of rotation of words of length $N$ are also separated by $C^{-4N}$.

Choosing $c$ small enough, we can guarantee that

$$C^{-4N} > C^{-4c\log s} = s^{-4c\log C} > s^{-1/8}$$

and thus each word of length $\leq N$ in the generators is a rotation through an angle $\geq s^{-1/8}$. This means that $d(x, g x) > 2s^{-1/4}$ for all points $x$ that are not close to a fixed point of $g$. But by the bound on the commutators, we deduce that any point $x$ can be close to the fixed point of at most one axis of rotation, and due to the condition that all non-trivial stabilizers are cyclic, this means that $x$ can be close to the fixed point of at most 2 words of length $N$, since any such words must be powers $g^n$ and $g^{-n}$ of the same element $g$.

**Proof of Theorem 1.4.** We keep the notation of the previous sections, and study the operator

$$W_{N, \alpha, s} = \sum_{n=1}^{N/2} \cos(2n\alpha) P_{2n}(T_q/2) \chi_s$$

where $\chi_s$ is the projection operator to the $s(s + 1)$-eigenspace with kernel $Z_s$, the averaging operator $T_q$ is defined as in Section 1.2, and $P_{2n}$ is the Chebyshev polynomials of the first kind of degree $2n$. Since $\chi_s$ is the projection to $\mathcal{H}_s$, any joint eigenfunction $\psi \in \mathcal{H}_s$ of $T_q$-eigenvalue $2\cos \alpha$ will satisfy $\chi_s \psi = \psi$ and thus by Lemma 2.2

$$\|\psi\|_p \lesssim \frac{1}{N} \|W_{N, \alpha, s} \psi\|_p \leq \frac{1}{N} \|W_{N, \alpha, s}\|_{L^2 \to L^p} \|\psi\|_2$$
and so it remains to show that
\[ \|W_{N,\alpha,s}\|_{L^2 \to L^p} \lesssim \sqrt{N} s^{\frac{1}{2} - \frac{2}{p}}, \]
which will follow from an estimate of the type
\[ \|S_n\chi_s\|_{L^p' \to L^p} \lesssim s^{1 - \frac{2}{p} q - \beta_n} \]
for some $\beta_p > 0$, by the arguments in Section 2. Naturally we will take $N \gtrsim p \log s$, in a way which will depend on $p$.

To prove (4.4), observe that the kernel $Z_s$ of $\chi_s$ is localized (in $L^{p/2}$) in the ball of radius $s^{-1/2}$ around the poles, since
\[ \|Z_s\|_{L^{p/2}} \lesssim \left( \int_{s^{-1/2}}^{\pi - s^{-1/2}} |Z_s(t)|^{p/2} \sin t \, dt \right)^{2/p} \]
\[ \lesssim \left( \int_{s^{-1/2}}^{\pi - s^{-1/2}} s^{p/4 t - p/4 + 1} \, dt \right)^{2/p} \]
\[ \lesssim s^{1/2 \left( t - \frac{4}{2} + 2 \frac{-s^{-1/2}}{s^{-1/2}} \right)^{2/p}} \]
\[ \lesssim s^{1/2 \left( s^{-1} - 1 \right)^{2/p}} \]
\[ \lesssim s^{3 - \frac{2}{p}}. \]
Note that for $p > 8$, this exponent $\frac{3}{4} - \frac{2}{p} < 1 - \frac{4}{p}$; i.e. it is of lower order than the $L^{p/2}$-norm of $Z_s$ inside the ball of radius $s^{-1/2}$, as computed in (4.3).

Recall now that
\[ \chi_s u(x) = Z_s u(x) = \int Z_s(d(x,y)) u(y) \, d\sigma(y). \]
We split $Z_s$ into three parts, $Z_s^{(1)} + Z_s^{(2)} + Z_s^{(3)}$, with $Z_s^{(1)}(\theta) = \eta_s(\theta) Z_s(\theta)$ and $Z_s^{(3)}(\theta) = \eta_s(\pi - \theta) Z_s(\theta)$, where $\eta_s \in C_0^\infty(\mathbb{R})$ is supported in $[0, s^{-1/4}]$ and is identically 1 on $[0, s^{-1/2}]$. We denote by $\chi_s = \chi_s^{(1)} + \chi_s^{(2)} + \chi_s^{(3)}$ the corresponding convolution operators. Recall that we are interested in estimating the norm of $S_n \chi_s = S_n \chi_s^{(1)} + S_n \chi_s^{(2)} + S_n \chi_s^{(3)}$. The kernel of $S_n \chi_s$ is given by
\[ [S_n \chi_s](x, y) = \frac{1}{q^n/2} \sum |g| = n Z_s(d(gx, y)), \]
so $S_n \chi_s$ is a normalized sum of convolutions with zonal spherical harmonics centered at different points.

We first treat the case of $S_n \chi_s^{(2)}$. Using Young’s inequality, we want to compute:
\[ \left( \int |S_n \chi_s^{(2)}(x, y)|^{p/2} \, d\sigma(x) \right)^{2/p} = \left( \int |S_n \chi_s^{(2)}(x, y)|^{p/2} \, d\sigma(y) \right)^{2/p}, \]
which is bounded by
\[ \frac{1}{q^n/2} \sum |g| = n \|Z_s^{(2)}\|_{p/2} \leq q^{n/2} \|Z_s^{(2)}\|_{p/2} \leq q^{n/2} s^{\frac{3}{4} - \frac{2}{p}}. \]
Now if \( n \leq 2\delta_p \log q s \) we obtain by Young’s inequality \(^2\)

\[
\|S_n\chi_s^{(2)}\|_{L^{p'} \to L^p} \lesssim s^{\frac{3}{2} - \frac{2}{p} + \delta_p}.
\]

(4.7)

We choose \( \delta_p \) so that \( \frac{3}{2} - \frac{2}{p} + \delta_p < 1 - \frac{4}{p} \), which is possible as long as \( p > 8 \).

By symmetry, the cases of \( S_n\chi_s^{(1)} \) and \( S_n\chi_s^{(3)} \) are identical. Here we use the fact that for any given \( y \in S^2 \) there are at most 2 words \( g \) of length \( n \) such that \( d(gx, y) \leq s^{-1/4} \), according to Lemma 4.2. This means we have at most 2 terms contributing to the sum

\[
[S_n\chi_s^{(1)}](x, y) = \frac{1}{q^{n/2}} \sum_{|g|=n} Z_s^{(1)}(d(gx, y))
\]

for each \( x, y \), and so up to a factor of 2 the kernel of \( S_n\chi_s^{(1)} \) is a normalized average over a disjoint union of copies of \( Z_s^{(1)} \)

\[
\left( \int [S_n\chi_s^{(1)}](x, y) |^{p/2} \, d\sigma(y) \right)^{2/p} \leq \frac{2^{p-2}}{q^{n/2}} \left( \sum_{|g|=n} \int |Z_s^{(1)}(d(gx, y))|^{p/2} \, d\sigma(y) \right)^{2/p} \leq 2q^{-n/2} \cdot q^{2n/p} \|Z_s^{(1)}\|_{p/2} \lesssim q^{n(\frac{4}{p}-2)} s^{1 - \frac{2}{p}}.
\]

(4.8)

By Young’s inequality we thus have

\[
\|S_n\chi_s^{(1)}\|_{L^{p'} \to L^p} \lesssim q^{-\frac{(p-4)}{p}} n s^{1 - \frac{4}{p}},
\]

(4.9)

which gives the necessary estimate\(^3\).

Putting (4.9) and (4.7) together and choosing \( \delta_p \) small enough we obtain

\[
\|S_n\chi_s\|_{L^{p'} \to L^p} \lesssim q^{-\beta_p n s^{1 - \frac{4}{p}}},
\]

for some \( \beta_p > 0 \). As in Lemma 2.5, this implies a similar bound

\[
\|P_{2n}(T_q/2)\|_{L^{p'} \to L^p} \lesssim q^{-\beta_p n s^{1 - \frac{4}{p}}}
\]

and the argument of the proof of Proposition 2.3 carries through to show

\[
\|W_{N,\alpha,s}\|_{L^2 \to L^p} \lesssim \sqrt{N} s^{\frac{1}{2} - \frac{2}{p}}
\]

as required.

\[\square\]

### 4.3. Outline of Sogge’s method for \( p > 6 \)

In this section we give a brief outline of Sogge’s method as it pertains to Theorem 1.4. Since the argument is more involved, and the application of our methods from Section 2 is no different from its application in the Young’s inequality argument above, we chose to present the latter, and will only sketch the relevant estimates from Sogge; for more details see [17, Theorem 5.1.1] (and the relevant arguments in the proof of Theorem 2.2.1 and Corollary 2.2.3). The main point is again that the main contribution to the norm of \( \chi_s\chi_s^* \) comes from

\(^2\)Here is where a direct appeal to Young’s inequality fails for \( 6 < p \leq 8 \), and one requires Sogge’s more refined methods to exploit the oscillations of the kernel; see Section 4.3 below.

\(^3\)In fact, this part of the estimate holds for all \( p > 4 \); it is only the bound on the \( S_n\chi_s^{(2)} \) term that determines the applicable values of \( p \).
points in a small $s^{-1+\delta}$-neighborhood, and since the images of these neighborhoods under $S_n$ are essentially disjoint, we get the desired estimates.

In order to show this concentration property in (4.5), we estimated the norm of the convolution kernel directly and applied a two-dimensional Young’s inequality. Though this gives the right estimate for $p > 8$, it is wasteful because it ignores the oscillations of the kernel. Sogge’s method sets up the integrals carefully in order to exploit these oscillations, which provide the necessary improvements in the estimates for $6 \leq p \leq 8$.

To begin with, Sogge replaces the spectral projector with an approximation $\tilde{\chi}_s$, such that by way of a parametrix the problem is reduced to studying a Fourier integral operator of the form

$$s^{-1/2} \tilde{\chi}_s f(x) = \int a_s(x, y) e^{is\psi(x,y)} f(y) d\sigma(y)$$

(4.10)

where the function $a_s$ and each of its derivatives is uniformly bounded independent of $s$, and the phase function $\psi$ satisfies certain non-degeneracy conditions [17, Lemma 5.1.3].

Using cutoff functions, we can divide $a_s$ into finitely many pieces, each compactly supported in a small neighborhood for convenience. We write

$$y = (y_1, r)$$

$$x = (x_1, t)$$

in (carefully chosen) local coordinates in each neighborhood, and consider for each $r$ the operator $T : L^2(\mathbb{R}) \to L^p(\mathbb{R}^2)$ given by

$$(Th)(x) = \int_{y_1} a_s(x; y_1, r) e^{is\psi(x; y_1, r)} h(y_1) dy_1$$

and in particular, fixing $r$ and writing $h(y_1) = f_r(y_1) = f(y_1, r)$, we examine

$$(T f)(x) = \int_{y_1} a_s(x; y_1, r) e^{is\psi(x; y_1, r)} f(y_1, r) dy_1$$

suppressing the dependence on $r$ and $s$. It is sufficient to obtain a bound on this $T$, since ([17, Corollary 2.2.3])

$$\| \int_{\mathbb{R}^2} a_s(\cdot; y_1, r) e^{is\psi(\cdot; y_1, r)} f(y_1, r) dy_1 dy_1 dr \|_{L^p(\mathbb{R}^2)}$$

(4.11)

$$\leq \int_r \| \int_{y_1} a_s(\cdot; y_1, r) e^{is\psi(\cdot; y_1, r)} f(y_1, r) dy_1 \|_{L^p(\mathbb{R}^2)} dr$$

$$\leq \int_r \| T \|_{L^2(\mathbb{R}) \to L^p(\mathbb{R}^2)} \cdot \| f(\cdot, r) \|_{L^2(\mathbb{R})} dr$$

$$\leq \| T \|_{L^2(\mathbb{R}) \to L^p(\mathbb{R}^2)} \cdot \| f \|_{L^2(\mathbb{R}^2)}$$

applying Hölder’s inequality in the last line, assuming as we may that the support of $f$ is contained in the compact support of $a_s$.

To study the operator $T$, we now freeze the variable $t$, and consider the operator $T_t$ given by the restriction $(T_t f)(x_1) = Tf(x_1, t)$, so that

$$(T_t f)(x_1) = \int_{y_1} a_s(x_1, t; y_1, r) e^{is\psi(x_1, t; y_1, r)} f(y_1, r) dy_1$$
The adjoint $T^*$ then decomposes as (writing $g_t(x_1) = g(x_1, t)$)

$$(T^*g)(y_1) = \int_{t \in \mathbb{R}} (T_t^*g_t)(y_1)dt$$

and therefore

$$(TT^*g)(x_1, t) = \int_{\mathbb{R}} (T_t T_{t'}^* g_{t'})(x_1)dt'$$

If the local coordinates are chosen properly to take advantage of the rapid oscillations of $e^{i\xi(t_1, t_2, y_1, y_2)}$, stationary phase approximations give the estimates \cite[(2.2.13)–(2.2.14)]{Sogge}

$$\|T_t T_{t'}^*\|_{L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})} \lesssim s^{-1} \quad (4.12)$$

$$|K_{t, t'}(x_1, x'_1)| \lesssim |s \cdot d((x_1, t); (x'_1, t'))|^{-1/2} \quad (4.13)$$

where $K_{t, t'}$ is the kernel of $T_t T_{t'}^*$. Note that $K_{t, t'}$ is uniformly bounded, so the estimate (4.13) is relevant for $|t - t'| > s^{-1}$; using the trivial bound for smaller $|t - t'| \leq s^{-1}$ naturally gives an estimate on the norm

$$\|T_t T_{t'}^*\|_{L^1 \rightarrow L^\infty} \lesssim \left\{ \begin{array}{ll}
1 & \text{if } |t - t'| \leq s^{-1} \\
s^{-1/2}|t - t'|^{-1/2} & \text{if } |t - t'| > s^{-1}
\end{array} \right. \quad (4.14)$$

Now, we have

$$\|(TT^*g)(\cdot, t)\|_{L^p(\mathbb{R})} = \left( \int \left( \int |(T_t T_{t'}^* g_{t'})(x_1)| dt' \right)^p dx_1 \right)^{1/p}$$

and since

$$\|TT^*g\|_{L^p(\mathbb{R}^2)} = \left( \int \left( \int |(TT^*g)(x_1, t)|^p dx_1 dt \right)^{1/p} \right)^{1/p}$$

we get

$$\|TT^*g\|_{L^p(\mathbb{R}^2)} \leq \left( \int \left[ \int \|T_t T_{t'}^*\|_{L^p \rightarrow L^p} \|g_{t'}\|_{L^p} dt' \right]^p dt \right)^{1/p}.$$ 

Using Young’s inequality we deduce that

$$\|TT^*g\|_{L^p(\mathbb{R}^2)} \leq C_{p/2} \left( \int \|g_{t'}\|_{L^p}^{p'} dt' \right)^{1/p'} = C_{p/2} \cdot \|g\|_{L^{p'}(\mathbb{R}^2)}$$

where

$$C_{p/2} = \max \left\{ \sup_t \left( \int \|T_t T_{t'}^*\|_{L^p \rightarrow L^p}^{p/2} dt' \right)^{2/p}, \sup_{t'} \left( \int \|T_t T_{t'}^*\|_{L^p \rightarrow L^p}^{p/2} dt \right)^{2/p} \right\}$$

For $p > 6$, we can use (4.14) to compute\footnote{Sogge uses the Hardy-Littlewood-Sobolev inequality in place of Young’s inequality to extend this to $p = 6$, but this is not relevant for our purposes here.}

$$C_{p/2} \lesssim s^{-2/p} \cdot \left( \int_{|t - t'| \leq s^{-1}} dt' \right)^{2/p} + s^{-1/2 - 1/p} \left( \int_{|t - t'| > s^{-1}} |t - t'|^{-\frac{p}{2} + \frac{1}{p}} dt' \right)^{2/p}$$

$$\lesssim s^{-2/p} \cdot s^{-\frac{3}{2}} + s^{-1/2 - 1/p} \cdot s^{\frac{1}{2} - \frac{3}{2} - \frac{3}{2} - \frac{3}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot s^{-4} \leq s^{-4} \quad (4.15)$$
and therefore we have
\[
\|TT^* g_t\|_{L^p(\mathbb{R})} \lesssim s^{-\frac{4}{p}} \|g_t\|_{L^{p'}(\mathbb{R})}
\] (4.16)
which in turn implies, by (4.11) and (4.10), the bound
\[
\|\tilde{\chi}_s \hat{X}\|_{L^{p'}(\mathbb{R}^2) \to L^{p'}(\mathbb{R}^2)} \lesssim s^{1 - \frac{4}{p}}
\]
as in Lemma 4.1, but now for all \(p > 6\).

Now in our setting we consider \(S_n \tilde{\chi}_s \hat{X}\), and wish to bound its norm from \(L^{p'} \to L^p\), so as in Section 4.2 we divide the kernel \(K_{t,t'} = K_{t,t'}^{(1)} + K_{t,t'}^{(2)}\) where \(K_{t,t'}^{(1)}\) is supported on \(d((x_1, t); (x'_1, t')) < s^{-1/4}\) and \(K_{t,t'}^{(2)}\) is supported on \(d((x_1, t); (x'_1, t')) > s^{-1/2}\). The bound (4.13) then gives
\[
\|K_{t,t'}^{(2)}\|_{\infty} \lesssim \begin{cases} 
  s^{-1/4} & |t - t'| \leq s^{-1/2} \\
  s^{-1/2} |t - t'|^{-1/2} & |t - t'| > s^{-1/2}
\end{cases}
\]
which gives via interpolation
\[
\|T_t T_{t'}^{*}\|_{L^{p'} \to L^p} \lesssim \begin{cases} 
  s^{-\frac{1}{2} - \frac{2}{p}} & |t - t'| \leq s^{-1/2} \\
  s^{-\frac{1}{2} - \frac{1}{p}} |t - t'|^{-\frac{1}{2}} & |t - t'| > s^{-1/2}
\end{cases}
\]
so that for \(p > 6\) we proceed as in (4.15), this time dividing the range of \(|t - t'|\) according to \(s^{-1/2}\), to estimate
\[
\|((TT^*)^{(2)} g(\cdot, t))\|_{L^p} \lesssim \left[ s^{-\frac{1}{2} - \frac{2}{p}} \left( \int_{\{|t' - t| \leq s^{-1/2}\}} dt' \right)^{2/p} + s^{-\frac{1}{2} - \frac{1}{2}} \left( \int_{\{|t' - t| > s^{-1/2}\}} |t - t'|^{-\frac{1}{2}} dt' \right)^{2/p} \right] \|g(\cdot, t)\|_{L^{p'}(\mathbb{R})}
\]
(4.17)
which as above means
\[
\|((TT^*)^{(2)})\|_{L^{p'} \to L^p} \lesssim s^{1 - \frac{5}{p}}
\]
and thus
\[
\|\tilde{\chi}_s \hat{X}\|_{L^{p'}(\mathbb{R}^2) \to L^p} \lesssim s^{\frac{2}{p}} = s^{1 - \frac{6}{p} - \delta_p}
\] (4.18)
for \(p > 6\), where \(\delta_p = \frac{1}{4} - \frac{2}{2p} > 0\).

This shows that the contribution of points at distance greater than \(s^{-1/2}\) is negligible for \(p > 6\), and in fact we can safely sum over copies of this part of the kernel to get
\[
\|S_n(\tilde{\chi}_s \hat{X}\|_{L^{p'} \to L^p} \lesssim s^{1 - \frac{4}{p} - \delta_p}
\]
if \(n\) and \(\delta_p\) are appropriately chosen, as above in (4.7). On the other hand, the remaining main part of the kernel is localized in disjoint balls of radius \(s^{-1/4}\) as in Section 4.2, and thus we may safely sum over these as well, gaining a power of \(q^{-\delta_n}\) in the one-dimensional Young’s inequality estimate for the integral (4.15), just as we obtained in (4.9). Thus the arguments of Section 4.2 carry through for \(p > 6\) as well, by using Sogge’s methods instead of the (wasteful) two-dimensional Young’s inequality.

It is worth mentioning that we cannot get an improvement here for \(p = 6\) itself—the reason is that in estimating Sogge’s integrals (4.17) for \(p = 6\), the contribution of

\footnote{Since we used cutoff functions in defining each \(a_s\), there is no need for a third piece covering the antipodal point.}
(TT∗)

(2)

is of the same order as (TT∗)

(1)

, and so the Sn sum over overlapping copies must be estimated trivially and we lose in the estimates. Our estimates for

p > 6 rely on the fact that (TT∗)

(2)

is of lower order than (TT∗)

(1)

, and so we have the necessary wiggle room of

δp

in the exponent (4.18) to accommodate summing over the overlapping copies. Accordingly, the implied constant of Theorem 1.4 will blow up as

p → 6.

References


Department of Mathematics, Bar-Ilan University, Ramat-Gan, 5290002 Israel

Email address: brookss@math.biu.ac.il

School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK

Email address: etienne.lemasson@bristol.ac.uk