WEIGHTED EIGENFUNCTION ESTIMATES WITH APPLICATIONS TO COMPRESSED SENSING

NICOLAS BURQ, SEMYON DYATLOV, RACHEL WARD, AND MACIEJ ZWORSKI

ABSTRACT. Using tools from semiclassical analysis, we give weighted L^∞ estimates for eigenfunctions of strictly convex surfaces of revolution. These estimates give rise to new sampling techniques and provide improved bounds on the number of samples necessary for recovering sparse eigenfunction expansions on surfaces of revolution. On the sphere, our estimates imply that any function can be efficiently recovered to within a factor of its best s-term approximation in the first N spherical harmonics from its values at $m \gtrsim sN^{1/6}\log^4(N)$ sampling points, improving on the previous bound of $m \gtrsim sN^{1/4}\log^4(N)$ necessary sampling points. In particular, any function having an s-sparse expansion can be recovered exactly from such undersampled measurements.

1. INTRODUCTION

Consider the sphere and a chosen rotational action generated by ∂_φ:

$$\mathbb{S}^2 := \{x \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = 1\}\,, \quad \partial_\varphi = x_1 \partial_{x_2} - x_2 \partial_{x_1}\,,$$

$$\mathbb{S}^2 \ni x = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta)\,, \quad 0 \le \theta \le \pi\,, \quad 0 \le \varphi \le 2\pi\,.$$

Let $k, \ell \in \mathbb{Z}$, $|k| \le \ell$, and let $Y_\ell^k(\varphi, \theta)$ be the L^2 normalized spherical harmonics, the joint eigenfunctions of the Laplacian in spherical coordinates $\Delta_{\mathbb{S}^2}$ and the rotational generator ∂_φ:

$$-\Delta_{\mathbb{S}^2} Y_\ell^k = -\left(\frac{1}{\sin^2\theta}\partial_\varphi^2 + \frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta)\right)Y_\ell^k = l(l+1)Y_\ell^k\,, \quad \frac{1}{i}\partial_\varphi Y_\ell^k = kY_\ell^k\,,$$

$$\int_0^{2\pi}\int_0^\pi Y_\ell^k(\varphi, \theta)\overline{Y_{\ell'}^{k'}(\varphi, \theta)}\sin\theta\,d\theta\,d\varphi = \delta_{\ell\ell'}\delta_{kk'}\,. \tag{1.1}$$

Applied to the sphere, our main result on weighted L^∞ estimates reads

Theorem 1. *Let $Y_\ell^k(\varphi, \theta)$, $0 \le \varphi \le 2\pi$, $0 \le \theta \le \pi$, be the spherical harmonics defined above. Then for $\ell \ge 1$,*

$$|\sin^2\theta\cos\theta|^{1/6}\,|Y_\ell^k(\varphi, \theta)| \le C\ell^{1/6}\,, \tag{1.2}$$

where C is a universal constant.

1

The power $\ell^{1/6}$ in (1.2) can be explained as follows. Taking the Fourier expansion in φ reduces the first differential equation in (1.1) to

$$\left(\frac{1}{\sin\theta}h\partial_\theta\left(\sin\theta h\partial_\theta\right) + \frac{\alpha^2}{\sin^2\theta} - 1\right)u = 0, \quad h = (\ell(\ell+1))^{-1/2}, \quad \alpha = hk,$$

where $k \in \mathbb{Z}$ is an eigenvalue of $\frac{1}{i}\partial_\varphi$ and $Y_\ell^k = u(\theta)e^{ik\varphi}$. When α is such that $0 < \epsilon < |\alpha| < 1 - \epsilon$, this equation has two turning points at $\sin\theta = \pm\alpha$. Physically this corresponds to caustic formation: the focusing at turning points increases the intensity of the wave function, that is, it increases the L^∞ norm by a factor of $h^{-1/6} \sim \ell^{1/6}$ according to Proposition 4.5 below.[1] Since the $h^{-1/6}$ loss happens all over the sphere, such growth in the L^∞ norm cannot be eliminated by a weight function. In order to get a uniform bound on the entire sphere in (1.2), we choose a weight function vanishing at the pole and the equator. A more detailed explanation of the weights and the principles of semiclassical analysis on which the analysis is based is given at the end of Section 3.

1.1. **Motivation.** Consider functions on the sphere which are *bandlimited* and *sparse*:

$$f(\varphi,\theta) = \sum_{\ell=0}^{\sqrt{N}-1}\sum_{k=-\ell}^{\ell} c_{\ell,k}Y_\ell^k(\varphi,\theta); \quad \text{at most } s < N \text{ of the } c_{\ell,k} \text{ are nonzero.} \quad (1.3)$$

Functions well-approximated as bandlimited and sparse arise in applications ranging from models for protein structure [22] to cosmic microwave background (CMB) data [1]. In [26], the authors showed that such functions can be efficiently reconstructed from far less information than their ambient dimension suggests; in particular, they showed that for certain sets of sampling points $(\varphi_i, \theta_i) \in \mathbb{S}^2$, $i \in [m]$, of size

$$m \gtrsim sN^{1/4}\log^4 N, \quad (1.4)$$

any function f of the form (1.3) can be reconstructed from its values $f(\varphi_i, \theta_i)$ as the function of bandwidth $\sqrt{N} - 1$ whose coefficient vector $c = c_{\ell,k}$ has minimal ℓ_1-norm $\|c\|_1 = \sum_{\ell=0}^{\sqrt{N}-1}\sum_{k=-\ell}^{\ell}|c_{\ell,k}|$. It is shown that m angular coordinates (φ_i, θ_i) where m satisfies (1.4), drawn independently from the measure $d\theta d\varphi$ on $[0,\pi] \times [0,2\pi]$, will almost always be a set of sampling points for which this holds.

Theorem 1 of this paper improves on the results in [26] by improving on the required number of sampling points for recovering functions of the form (1.3) to

$$m \gtrsim sN^{1/6}\log^4(N) \quad (1.5)$$

by drawing angular coordinates (φ_i, θ_i) independently from the measure $|\tan(\theta)|^{1/3}d\theta d\varphi$ on $[0,\pi] \times [0,2\pi]$. The specific statement is given in Corollary 1. As seen in Figure 1.1(c), this measure generates higher sampling density around the poles and equator of the sphere; the

---

[1]The $h^{-1/6}$ factor can be seen on the model example of the equation $(h^2D_x^2 + x)v = 0$, with a turning point at $x = 0$ and locally L^2-normalized solution $h^{-1/6}Ai(h^{-2/3}x)$, where Ai is the Airy function. The $h^{-1/6}$ normalization here follows from the asymptotic behavior of $Ai(y)$ as $y \to -\infty$.

measure $d\theta d\varphi$, illustrated in Figure 1.1(b) and on which the analysis of [26] is based, only generates higher sampling density at the poles.

It remains open whether the factor of $N^{1/6}$ in (1.5) can be eliminated. As the discussion following Theorem 1 indicates, such a result cannot be proven using weight functions alone.

1.2. **Sparse recovery for arbitrary surfaces of revolution.** The weighted L^∞ estimates given in Corollary 3 of Section 3 provide sampling strategies more broadly for recovering sparse eigenfunction expansions on any strictly convex surface of revolution. In particular, assume that M is a strictly convex surface of revolution parametrized by $(r, \varphi) \in [r_-, r_+] \times [0, 2\pi)$. The induced Riemannian metric on M is given by

$$g = dr^2 + a(r)^2 d\varphi^2, \quad a(r) = (r_+ - r)(r - r_-)b(r), \quad b(r_\pm) > 0, \quad b'(r_\pm) = 0,$$

where $a(r)$ has a unique nondegerate local maximum at $r = r_0$, $r_- < r_0 < r_+$: $a'(r) \neq 0$, $r \neq r_0$, $a''(r_0) < 0$. In particular, using Corollary 3 in Section 3, we will prove the following.

Proposition 1.1. *Suppose that M is a strictly convex surface of revolution and consider $\{\psi_j\}$, the (L^2-normalized) joint eigenfunctions of the Laplace-Beltrami operator on M and the rotational generator $\frac{1}{i}\partial_\varphi$.*

Let m, s, and N be given integers satisfying

$$m \gtrsim sN^{1/6}\log^4(N), \tag{1.6}$$

and suppose that m coordinates $(\varphi_i, r_i)_{i=1}^m$ are drawn independently from the measure

$$\left(\frac{a(r)}{|r - r_0|}\right)^{1/3} dr d\varphi$$

on $[r_-, r_+] \times [0, 2\pi)$. With probability exceeding $1 - N^{-\log^3(s)}$ with respect to the draw of the coordinates, any function $f(\varphi, r) = \sum_{j=1}^N c_j \psi_j(\varphi, r)$ which is s-sparse, $\|c\|_0 \leq s$, is recovered exactly from the samples $\left(f(\varphi_i, r_i)\right)_{i=1}^m$ as the minimizer of an ℓ_1-minimization program:

$$c = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad \sum_{j=1}^N z_j \psi_j(\varphi_i, r_i) = f(\varphi_i, r_i), \quad i \in [m]. \tag{1.7}$$

Applying Proposition 1.1 to the sphere, we get the following corollary.

Corollary 1. *With high probability with respect to the draw of $m \gtrsim sN^{1/6}\log^4 N$ i.i.d. sampling points on the sphere (φ_i, θ_i) from the measure $|\tan^{1/3}(\theta)|d\theta d\varphi$ on $[0, \pi] \times [0, 2\pi]$, any s-sparse function of the form (1.3) is recovered exactly as the minimizer of the convex program (1.7).*
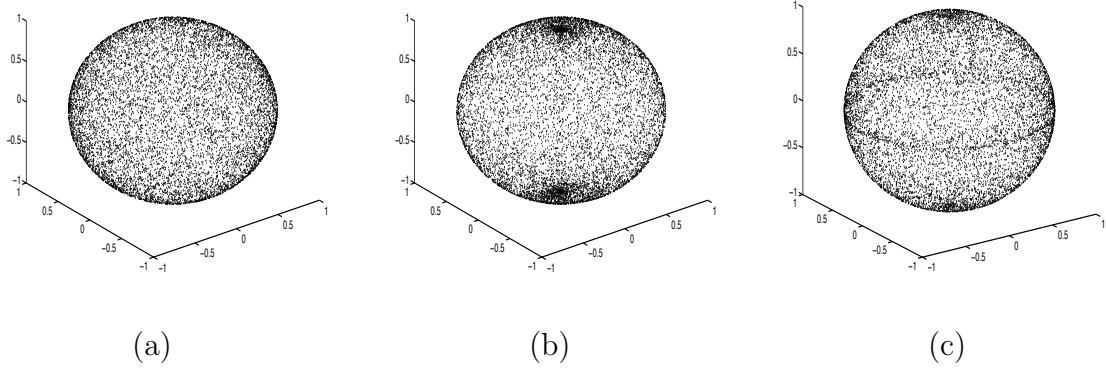
FIGURE 1.1. The spherical sampling measures (a) $\sin(\theta)\,d\theta d\varphi$, (b) $d\theta d\varphi$, and (c) $|\tan(\theta)|^{1/3}\,d\theta d\varphi$, as illustrated through $m = 10,000$ i.i.d. samples from each measure.

1.3. **Numerical experiments.** In this section we assess the impact of the sparse recovery results in Corollary 1 for problem sizes of practical interest. We draw m sampling points (φ_i, θ_i) i.i.d. from each of the measures (a) $\sin(\theta)\,d\theta d\varphi$, (b) $d\theta d\varphi$, and (c) $|\tan(\theta)|^{1/3}\,d\theta d\varphi$, as illustrated in Figure 1.1, and compare the rate of correct reconstruction of sparse bandlimited functions via ℓ_1-minimization (1.7) for each of the three sampling schemes. More specifically, for each choice of sampling measure, we vary a number of sampling points m between 1 and N, and vary a sparsity level s between 1 and m. For each choice of m and s, we generate 50 s-sparse bandlimited functions by repeatedly choosing a support $S \subset [N]$ of size $|S| = s$ at random, and prescribing to the chosen support i.i.d. Gaussian coefficients.

From left to right, the phase diagrams in Figure 1.2 correspond to sampling measures (a) $\sin(\theta)d\theta d\varphi$, (b) $d\theta d\varphi$, and (c) $|\tan(\theta)|^{1/3}d\theta d\varphi$. White indicates complete recovery, and black indicates no recovery whatsoever. It is clear that the sampling strategies (b) and (c) give better results than (a). Diagrams (b) and (c) both show a sharp transition between complete recovery and no recovery whatsoever as the ratio s/m increases as a function of m/N. However, the region of phase space corresponding to complete recovery is noticeably larger in (3) when m/N is large. Note that when $m = N$, all three sampling schemes should give perfect reconstruction as the system of equations $y = Ac$ in the minimization problem (1.7) has a unique solution with probability 1. However, plots (a) and (b) show zero reconstruction around this point, an artifact of round-off error due to the ill-conditioning of the corresponding sampling schemes.

Organization of the paper. In Section 2 we review the relationship between sparse recovery techniques on manifolds and weighted L^∞ bounds on the associated eigenfunctions. We then show how the main results of this paper strengthen and generalize existing sparse
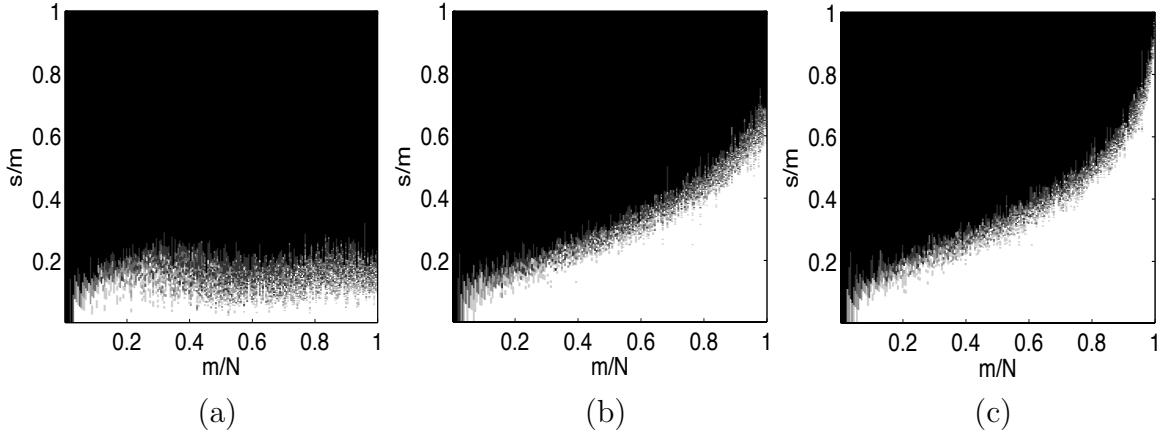
(a)   (b)   (c)

FIGURE 1.2. Phase diagrams illustrating transition between uniform recovery (white) to no recovery whatsoever (black) of sparse spherical harmonic expansions $f(\varphi, \theta) = \sum_{\ell=0}^{19} \sum_{k=-\ell}^{\ell} c_{\ell,k} Y_{\ell}^{k}(\varphi, \theta)$, $\|c\|_0 \le s$, from samples $f(\varphi_i, \theta_i)$. In (a), the sampling points are drawn from the volume measure $\sin \theta \, d\theta d\varphi$ illustrated in Figure 1.1(a). In (b) the sampling points are drawn from the Cartesian measure $d\theta d\varphi$ illustrated in Figure 1.1(b), and in (c) the sampling points are drawn from the measure $|\tan(\theta)|^{1/3} \, d\theta d\varphi$ illustrated in Figure 1.1(c).

recovery bounds. The generalization of Theorem 1 to arbitrary convex surfaces of revolution is given in Theorem 2 of Section 3, while Section 4 provides a detailed account of preliminaries from semiclassical analysis needed for the proof which is presented in Section 5.

Notation. In the paper C denotes a constant, independent of asymptotic parameters, but changing depending on the context. We use the usual $\mathcal{O}$ notation with subscripts to indicate that the associated constant might depend on the variable in the subscript, for instance $f = \mathcal{O}_x(g)$ means that $f(x, y) \le C(x)g(y)$ for some $C(x)$ depending on x. We follow the basic notational convention listed in [31, Appendix A]. Consequently the above notation should not be confused with $u = \mathcal{O}_V(g)$ for V a Hilbert space; the latter means that $\|u\|_V \le Cg$. The notation $f \lesssim g$ means that there exists C such that $f \le Cg$. Finally, we use the shorthand $[N] = \{1, 2, ..., N\}$. A vector $x \in \mathbb{C}^N$ or $x \in \mathbb{R}^N$ is called *s-sparse* if it is supported on an index set $S \subset [N]$ of cardinality $|S| \le s$, and we specify an s-sparse vector using the notation $\|x\|_0 := |S| = \{\#j : |x_j| > 0\}$.

2. COMPRESSED SENSING AND WEIGHTED L^∞ ESTIMATES

Suppose we have a finite system of functions $\{\psi_j : j \in [N]\}$ on a compact manifold M. Suppose we also have a function $f : M \to \mathbb{C}$ which is s-sparse with respect to this function

system,

$$f = \sum_{j=1}^{N} c_j \psi_j, \quad \|c\|_0 \leq s < N. \tag{2.1}$$

The area of compressed sensing [4] is concerned with the following questions. For a given system $\{\psi_j\}$ and s-sparse function f of the form (2.1), how many samples $f(x_i)$ where $x_i \in M$ do we need to uniquely identify f? Is it possible moreover to efficiently and stably reconstruct such a function from these samples? To distinguish a linear combination of the N functions ψ_j we would clearly need N samples. On the other hand, if f is s-sparse and the locations of its s nonzero coefficients c_j are known, then we would need only s samples to distinguish f. If we know a priori that f is s-sparse but we do not know the locations of the s coefficients, then $2s$ samples can still suffice for distinguishing f. Consider the matrix $\Psi \in \mathbb{C}^{m \times N}$ with entries $\Psi_{i,j} = \psi_j(x_i)$, and observe that a vector of m samples can be written as

$$\big(f(x_1), f(x_2), \ldots, f(x_m)\big)^t = \Psi c. \tag{2.2}$$

Each s-sparse function f has a distinct image $y = \Psi c$ if every sub-matrix of the $m \times N$ matrix Ψ consisting of at most $2s$ columns is non-singular, and this is true for many $m \times N$ matrices having only $m = 2s$ rows (consider matrices having i.i.d. Gaussian entries, for example.) Subject to this "restricted invertibility" condition, one could distinguish the unique s-sparse solution to $y = \Psi c$ by minimizing over all s-sparse vectors c. In general, this is an NP-hard problem. However, if we strengthen the "restricted invertibility" condition to a certain restricted *isometry* condition — a property that can only hold if Ψ has at least $m \gtrsim s \log(N/s)$ rows [7], [13], [9] — then polynomial-time sparse recovery algorithms are possible. In the compressed sensing literature, a matrix $\Psi \in \mathbb{C}^{m \times N}$ is said to have the *restricted isometry property* of order $2s$ if, for a fixed parameter $\delta < 1$,

$$(1 - \delta)\|u\|_2 \leq \|\Psi u\|_2 \leq (1 + \delta)\|u\|_2, \quad \forall u : \|u\|_0 \leq 2s. \tag{2.3}$$

As shown in [4], if a matrix $\Psi \in \mathbb{C}^{m \times N}$ has this property, and if $y = \Psi c$ for some s-sparse vector $c \in \mathbb{C}^N$, then c is guaranteed to also be the vector of minimal ℓ_1-norm among solutions z to the underdetermined system $\Psi z = y$. That is,

$$c = \arg\min_z \|z\|_1 \quad \text{subject to } \Psi z = \Psi c. \tag{2.4}$$

Moreover for an arbitrary vector $c \in \mathbb{C}^N$, the solution $c^{\#}$ to the ℓ_1-minimization program (2.4) will satisfy the error bound

$$\|c - c^{\#}\|_2 \lesssim \|c - c_0\|_1 / \sqrt{s}, \tag{2.5}$$

where c_0 is the *best s-sparse approximation* to c, satisfying $\|c - c_0\|_p = \arg\min_{z \in \mathbb{C}^N : |z| \leq s} \|c - z\|_p$, and being the s-sparse vector which is equal to c over the support of the s largest-magnitude entries of c and zero everywhere else.

Recall that ℓ_1-minimization can be solved efficiently as a linear programming problem in the real-valued case and as a second-order cone programming problem in the general case.

2.1. Sparse recovery for bounded orthonormal systems.
In general the restricted isometry property (2.3) is hard to verify for a given matrix Ψ. For certain random matrices, this property can nevertheless be guaranteed with high probability. Suppose we have a system of functions $\{\psi_j : j \in [N]\}$ which are orthonormal on a measurable space M endowed with a probability measure ν, i.e.

$$\int_M \psi_j(x)\overline{\psi_i(x)}d\nu(x) = \delta_{i,j}, \quad i,j \in [N]. \tag{2.6}$$

Suppose further that m sampling points $x_i \in M$ are drawn independently from the orthogonalization measure ν. Then, as shown in [23], with high probability with respect to the draw of the sampling points, the normalized sampling matrix $\frac{1}{\sqrt{m}}\Psi$, where $\Psi_{i,j} = \psi_j(x_i)$, satisfies the restricted isometry property (2.3) as long as the number of samples $m \gtrsim B^2 s \log^4(N)$, where

$$B = \max_j \|\psi_j\|_\infty. \tag{2.7}$$

The parameter B should be interpreted as a measure of *incoherence* between the basis ψ_j and pointwise measurements; the smaller B, the fewer number m of sampling points are needed to recover sparse expansions (2.1). This can be interpreted as a discrete Heisenberg uncertainty principle [10]. A precise statement follows.

Proposition 2.1. *Suppose $\{x_i : i \in [m]\}$ is a set of independent and identically distributed (i.i.d.) sampling points drawn from the orthogonalization measure ν associated to an orthonormal system of functions $\{\psi_j : j \in [N]\}$ with uniform bound $B = \max_j \|\psi_j\|_\infty$. If*

$$m \gtrsim B^2 s \log^4(N), \tag{2.8}$$

then with probability at least $1 - N^{-\log^3(s)}$, the following holds for all $f(x) = \sum_{j=1}^N c_j \psi_j(x)$. From samples $(f(x_i))_{i=1}^m$, the minimizer

$$c^\# = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \text{ subject to } \sum_{j=1}^N z_j \psi_j(x_i) = f(x_i), \quad i \in [m] \tag{2.9}$$

satisfies the error bound

$$\|c - c^\#\|_2 \lesssim \|c - c_0\|_1/\sqrt{s}, \tag{2.10}$$

where c_0 is the best s-sparse approximation to c. If f is s-sparse, then reconstruction is exact $c^\# = c$.

Let us apply Proposition 2.1 to a concrete example. The orthonormal system of complex exponentials $\psi_j(t) = e^{2\pi ijt}, t \in [0,1]$, has optimal uniform bound $B = \max_j \|\psi_j\|_\infty = 1$. Applying Proposition 2.1, we see that $m \gtrsim s \log^4(N)$ sampling points drawn independently

from the uniform measure on $[0, 1]$ will be sufficient to identify any s-sparse trigonometric polynomial of degree at most N.

The complex exponentials are eigenfunctions of the Laplacian on the circle. More generally, for a compact n-dimensional Riemannian manifold, the L^2-normalized eigenfunctions with eigenvalue λ are bounded in L^∞ by $\lambda^{(n-1)/4}$ – see [31, Section 7.4] and references given there. Since the number of eigenvalues less than λ behaves like $N = \lambda^{n/2}$ (see [17] or [31, Section 14.3]), we obtain a uniform bound on the first N eigenfunctions of a general n-dimensional manifold:

$$B \simeq N^{\frac{n-1}{2n}}.$$

Applying Proposition 2.1 gives the following.

Corollary 2. *Let (M, g) be a compact n-dimensional Riemannian manifold and let $\{\psi_j : j \in [N]\}$ be the first N eigenfunctions of the Laplacian on M (with respect to the ordering of eigenfunctions).*

Suppose $\{x_i : i \in [m]\}$ is a set of independent and identically distributed sampling points drawn from the measure given by the Riemannian volume. If the number of sampling points satisfies

$$m \gtrsim sN^{\frac{n-1}{n}} \log^4(N), \tag{2.11}$$

then with probability at least $1 - N^{-\log^3(s)}$, the following holds for all $f(x) = \sum_{j=1}^N c_j \psi_j(x)$.

If Ψ is the sampling matrix associated to the eigenfunctions and sampling points, and if $c^\#$ is the ℓ_1-minimizer as defined in (2.9), then

$$\|c - c^\#\|_2 \lesssim \|c - c_0\|_1 / \sqrt{s},$$

where c_0 is the s-term approximation to c.

When $n = 1$, we recover the near-optimal number of measurements $m \gtrsim s \log^4(N)$. However, even for $n = 2$, the number of measurements $m \gtrsim sN^{1/2} \log^4 N$ represents is significantly weaker result, and as the dimension n increases, the bound (2.11) becomes weaker and weaker as the basis functions can become more and more localized on the manifold. In [25], Proposition 2.1 was generalized to give estimates with a wider range of applicability.

Proposition 2.2 (From [25]). *Let $\{\psi_j : j \in [N]\}$ be an orthonormal system of functions on a probability space M with orthogonalization measure ν.*

Suppose that $\omega : M \to \mathbb{R}$ satisfies $\int_M \omega(x)\nu(x)dx = 1$, and suppose that the functions $Q_j(x) = \omega(x)^{-1/2}\psi_j(x)$ are bounded:

$$\sup_{j \in [N]} \sup_{x \in M} |Q_j(x)| \le K. \tag{2.12}$$

Suppose that $m \gtrsim K^2 s \log^4(N)$, and suppose that $\{x_i : i \in [m]\}$ are i.i.d. sampling points from the composite orthogonalization measure $\mu = \omega\nu$.

Then with probability at least $1 - N^{-\log^3(s)}$ *the following holds for all* $f(x) = \sum_{j=1}^{N} c_j \psi_j(x)$.
The minimizer

$$c^{\#} = \arg \min_{z \in \mathbb{C}^N} \|z\|_1 \text{ subject to } \sum_{j=1}^{N} c_j \psi_j(x_i) = f(x_i), \quad i \in [m], \qquad (2.13)$$

satisfies the error bound

$$\|c - c^{\#}\|_2 \lesssim \|c - c_0\|_1 / \sqrt{s}, \qquad (2.14)$$

where c_0 *is the best s-sparse approximation to c.*

Proof. The system $\{Q_j\}_{j \in [N]}$, where $Q_j(x) = \omega(x)^{-1/2} \psi_j(x)$, satisfies $\sup_{j \in [N]} \|Q_j\|_\infty \leq K$. Moreover, by design the Q_j are orthonormal with respect to the composite measure $\mu = \omega \nu$. Noting that the set of linear constraints $\sum_{j=1}^{N} c_j \psi_j(x_i) = f(x_i)$ are equivalent to the constraints $\omega(x_i)^{-1/2} \sum_{j=1}^{N} c_j \psi_j(x_i) = \omega(x_i)^{-1/2} f(x_i)$, we apply Proposition 2.1 to the system $\{Q_j\}$ to derive the stated result. $\qquad\square$

Remark: In Proposition 2.1, and in turn Proposition 2.2 and Proposition 1.1, the ℓ_1-minimization program can be adapted to be robust to *measurement noise* $y_i = f(x_i) \pm \eta_j$; for details, we refer the reader to [23].

Proposition 2.2 elucidates the connection between weighted L^∞ estimates on orthonormal function systems and sparse recovery guarantees. Below we summarize how Proposition 2.2, applied to the weighted L^∞ estimates of this paper, improves and generalizes previous results on sparse function recovery on the sphere and gives rise to Proposition 1.1.

(1) In [20], Krasikov proved the following weighted L^∞ estimate for the spherical harmonics Y_ℓ^k:

$$(\sin \theta)^{1/2} |Y_\ell^k(\varphi, \theta)| \lesssim \ell^{1/4}.$$

This estimate implies

$$\sup_{0 \leq \ell \leq \sqrt{N}-1} \sup_{-\ell \leq k \leq \ell} (\sin \theta)^{1/2} |Y_\ell^k(\varphi, \theta)| \lesssim N^{1/8}. \qquad (2.15)$$

In [26], the authors applied Proposition 2.2 using this estimate to conclude that $m \gtrsim s N^{1/4} \log^4 N$ sampling points on the sphere with angular coordinates (φ_i, θ_i) drawn independently from the measure $d\theta d\varphi$ on $[0, \pi] \times [0, 2\pi)$ suffice for recovering sparse spherical harmonic expansions of the form (1.3). This improves on the $m \gtrsim s N^{1/2} \log^4(N)$ required sampling points drawn i.i.d. from the uniform distribution on the sphere as given by Corollary 2.

(2) The weighted L^∞ estimate

$$|\sin^2 \theta \cos \theta|^{1/6} |Y_\ell^k(\varphi, \theta)| \leq C \ell^{1/6} \qquad (2.16)$$

given in Theorem 1 provides even stronger sparse recovery guarantees for spherical harmonic expansions. Applying this estimate to Proposition 2.2 gives Corollary

1: $m \gtrsim sN^{1/6}\log^4(N)$ sampling points (φ_i, θ_i) from the measure $|\tan(\theta)|^{1/3}d\theta d\varphi$ suffice for recovering sparse spherical harmonic expansions (1.3).

(3) As summarized in Proposition 1.1, the weighted L^∞ estimate of Corollary 3 gives rise to more general sampling strategies for recovering sparse eigenfunction expansions on strictly convex surfaces of revolution.

3. Weighted eigenfunction estimates for surfaces of revolution

If M be a smooth surface of revolution, let ∂_φ be the vector field generating the action of the circle $\mathbb{S}^1 = \mathbb{R}/(2\pi\mathbb{Z})$ on M by rotations around the axis of revolution. Denote by Δ the Laplace–Beltrami operator on M, and by D_φ the self-adjoint operator $\frac{1}{i}\partial_\varphi$ which commutes with Δ. This follows the standard convention for the operators quantizing momenta.

Let $h > 0$ be a small parameter, and assume that $u \in \mathcal{C}^\infty(M)$ satisfies the conditions

$$\|u\|_{L^2} \leq C_0\,, \tag{3.1}$$

$$\|(-h^2\Delta - 1)u\|_{L^2} \leq C_0 h\,, \tag{3.2}$$

$$(hD_\varphi - \alpha)u = 0\,. \tag{3.3}$$

Here $\alpha \in h\mathbb{Z}$ varies in a fixed compact set and C_0 is some fixed constant. Both $-h^2\Delta - 1$ and $hD_\varphi - \alpha$ are semiclassical differential operators; we will freely use the notation of semiclassical analysis that can be found, for example, in [31, Chapter 4] (see also [21]).

We also assume that u satisfies the following localization assumption: there exists a compactly microlocalized operator $X(h)$ (that is, $X(h) = \psi^w(x, hD_x) + \mathcal{O}_{H_h^{-N}\to H_h^N}(h^N)$ for some $\psi \in \mathcal{C}_c^\infty(T^*M)$ and each N) and fixed constants C_N such that for each N,

$$\|(1 - X(h))u\|_{H_h^N} \leq C_N h^N. \tag{3.4}$$

Remark: Conditions (3.1), (3.2), and (3.4) are in particular satisfied if u is an L^2 normalized eigenfunction of $-\Delta$ for an eigenvalue in the segment

$$h^{-2}[1 - C_0 h, 1 + C_0 h]\,,$$

as applied in Proposition 1.1.

The weaker condition (3.2) has the advantage that it is local:

Proposition 3.1. *Assume that $\chi \in \mathcal{C}^\infty(M)$ and $D_\varphi\chi = 0$. If u satisfies* (3.1)–(3.3), *then χu satisfies these conditions as well, possibly with larger value of the constant C_0.*

Similarly, if condition (3.4) *holds for u, it holds for χu.*

Proof. Conditions (3.1) and (3.3) for χu are trivially satisfied; we now verify (3.2). Since the commutator $[-h^2\Delta, \chi]$ is equal to h times a semiclassical differential operator of order
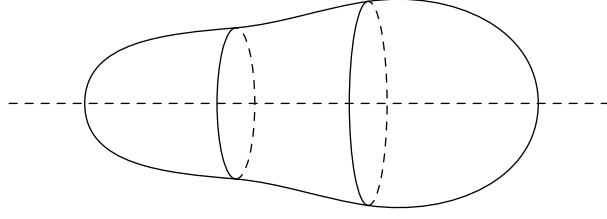
FIGURE 3.1. A surface of revolution

1, we have

$$\|(-h^2\Delta - 1)\chi u\|_{L^2} \leq \|\chi(-h^2\Delta - 1)u\|_{L^2} + \|[-h^2\Delta, \chi]u\|_{L^2}$$
$$= \mathcal{O}(h(1 + \|u\|_{H_h^1})).$$

However,

$$\|u\|_{H_h^1}^2 \sim ((-h^2\Delta + 1)u, u)_{L^2} = \mathcal{O}(1)$$

by (3.1) and (3.2). Here H_h^s denotes the semiclassical Sobolev space defined using the norm $\|(I - h^2\Delta_M)^{s/2}u\|_{L^2}$.

To verify (3.4), we use that $\chi u = \chi X(h)u + \mathcal{O}_{\mathcal{C}^\infty}(h^\infty)$; however, if $Y(h)$ is a compactly microlocalized pseudodifferential operator equal to the identity microlocally near the wave-front set of $X(h)$ (and thus of $\chi X(h)$), then $(1-Y(h))\chi u = (1-Y(h))\chi X(h)u + \mathcal{O}_{\mathcal{C}^\infty}(h^\infty) = \mathcal{O}_{\mathcal{C}^\infty}(h^\infty)$.

$\square$

Proposition 3.1 implies that, if we want to obtain weighted L^∞ (or any other local) estimates on every function u satisfying (3.1)–(3.3), it is enough to cover M by open sets invariant under rotation (which we will call bands) and prove the estimates for functions supported in each of these bands. The next result provides weighted L^∞ estimates for three common types of behavior of the metric in bands:

Theorem 2. *Let $U_\varepsilon \subset M$ be a band given by one of the three cases below; the small parameter $\varepsilon > 0$ characterizes the width of this band. Then for $\varepsilon > 0$ and $h > 0$ small enough and some constant C, and for each $\alpha \in h\mathbb{Z}$ varying in a fixed compact set, each function $u \in \mathcal{C}^\infty(M)$ supported in U_ε and satisfying (3.1)–(3.4) has the following weighted L^∞ estimates:*

(1) *Regular case: U_ε has coordinates $(r, \varphi) \in (-\varepsilon, \varepsilon) \times \mathbb{S}^1$ and the metric*

$$g = g(r)[dr^2 + d\varphi^2]. \tag{3.5}$$

Here g is a smooth function independent of ε, and $g(0) > 0$, $g'(0) \neq 0$. The corresponding estimate is

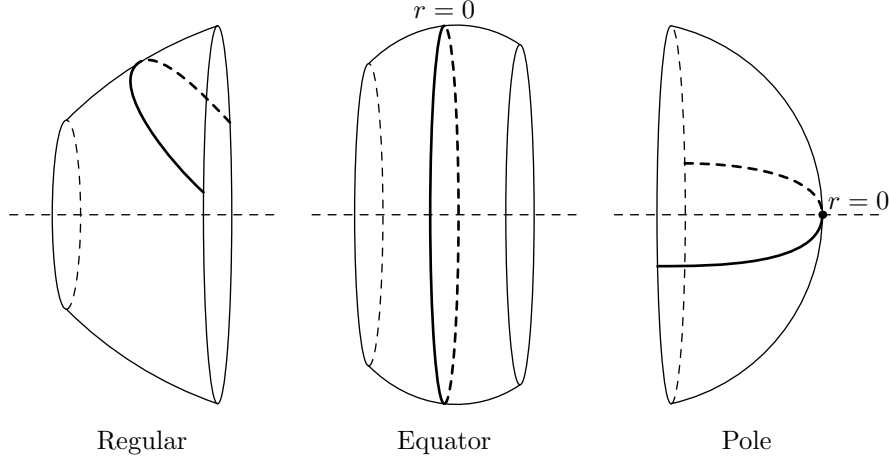$$|u(r, \varphi)| \leq Ch^{-1/6}. \tag{3.6}$$

FIGURE 3.2. Three considered types of bands, with a geodesic shown for each type

(2) *Elliptic equator: U_ε has the same coordinates and metric as in the regular case, but $g(0) > 0$, $g'(0) = 0$, and $g''(0) < 0$. The corresponding estimates are*

$$|u(r, \varphi)| \leq C \min(h^{-1/4}, h^{-1/6}r^{-1/6}, h^{-1/6}|g(0) - \alpha^2|^{-1/12}); \tag{3.7}$$

$$|u(r, \varphi)| \leq Cr^{-1/2} \quad for \ |g(0) - \alpha^2| < \varepsilon r^2. \tag{3.8}$$

(3) *Pole: U_ε has coordinates $(x, y) = (r \cos \varphi, r \sin \varphi)$, with $|x|^2 + |y|^2 < \varepsilon^2$, and the metric*

$$g = g(x^2 + y^2)[dx^2 + dy^2] = g(r^2)[dr^2 + r^2 \, d\varphi^2], \tag{3.9}$$

where g is a smooth function and $g(0) > 0$. The corresponding estimates are

$$|u(r, \varphi)| \leq C \min(h^{-1/2}, h^{-1/6}r^{-1/3}, h^{-1/6}|\alpha|^{-1/3}); \tag{3.10}$$

$$|u(r, \varphi)| \leq Cr^{-1/2} \quad for \ |\alpha| < \varepsilon r. \tag{3.11}$$

Remark. The metric of a surface of revolution can be brought locally to the form (3.5) or (3.9), with $g(0) > 0$, with no conditions on the derivatives of g. Indeed, away from the poles (points on the surface lying on the axis of rotation), the metric has the form

$$(1 + \tilde{g}'(\tilde{r})^2)d\tilde{r}^2 + \tilde{g}(\tilde{r})^2 \, d\varphi^2,$$

where $\tilde{r}$ is the projection onto the axis of rotation and $\tilde{g}$ is a positive function giving the profile of the surface. Making a change of variables $\tilde{r} \to r$ with $dr/d\tilde{r} = \tilde{g}(\tilde{r})^{-1}\sqrt{1 + \tilde{g}'(\tilde{r})^2}$, we bring the metric to the form (3.5). The case of a pole is handled similarly. As $(\tilde{r}, \varphi)$ we can use the geodesic polar coordinates with respect to the pole and a different change of variables $\tilde{r} \to r$.

We obtain the following corollary for convex surfaces of revolution. The estimate is the analogue of the estimate (1.2) in the case of the sphere:

Corollary 3. *Suppose that M is a strictly convex surface of revolution parametrized by $(r, \varphi) \in [r_-, r_+] \times [0, 2\pi)$, with the metric $dr^2 + a(r)\, d\varphi^2$ as in the discussion preceeding Proposition 1.1. Suppose that*

$$(\Delta - \lambda)u = 0, \quad (\tfrac{1}{i}\partial_\varphi - k)u = 0, \ \|u\|_{L^2(M)} = 1. \tag{3.12}$$

Then,

$$a(r)^{1/3}|r - r_0|^{1/6}|u(r, \varphi)| \leq C\lambda^{1/12}. \tag{3.13}$$

The Riemannian volume measure on M is given by $a(r)drd\varphi$ which means that the sampling measure based on (3.13) should be given by

$$\left(\frac{a(r)}{|r - r_0|}\right)^{1/3} drd\theta;$$

here $a(r)/(r - r_0)$ is the replacement for $\tan\theta$. The constant K in (2.12) is given by

$$K \simeq N^{1/12}.$$

Theorem 2 does not cover the case of a band with metric of the form (3.5) and $g'(0) = 0$, $g''(0) \leq 0$; in particular, it does not apply to the case of a hyperbolic equator, when $g''(0) < 0$. This case does not occur for convex surfaces considered here.

Let us give an informal explanation of the estimates in Theorem 2. By considering an eigenfunction decomposition of u, we can reduce to the case when u is an exact joint eigenfunction of $-h^2\Delta$ and hD_φ, rather than a function satisfying (3.2). Using semiclassical analysis (see [31, 21] for the general theory and the references below for specific facts we will be using), we can relate the behavior of the 'quantum' object u for small values of the 'Planck constant' h to the corresponding 'classical' integrable Hamiltonian system given by the principal symbols p and q of $-h^2\Delta$ and hD_φ, respectively. The principal symbol of a differential operator is a polynomial (on each fiber) function on the cotangent bundle T^*M, obtained formally by replacing each instance of hD_{x_j} by the corresponding momentum ξ_j, and then discarding the terms of higher order in h. In our situation, p is the square of the norm induced by g on the cotangent bundle and q is the momentum corresponding to φ. The function u will then be concentrated, or microlocalized, on the set

$$\Lambda = \{p = 1, \ q = \alpha\} \subset T^*M.$$

One can in fact approximate u by certain explicit highly oscillating integral expressions up to an $\mathcal{O}(h^\infty)$ error; our analysis would consist of studying the asymptotic behavior of these integrals as $h \to 0$. The set Λ consists of unit geodesics with prescribed angular momentum (that is, of rotations of one such geodesic); there are two possibilities:

(1) Λ is a Lagrangian torus;
(2) Λ is a circle corresponding to an equator.

In case (1), u is a Lagrangian distribution associated to Λ – that is, it can be written as a finite sum of expressions of the WKB form (4.1), with $\Phi(x, \theta)$ locally parametrizing the Lagrangian Λ and a some smooth symbol – see Section 4.1 for details. The L^∞ norm of u corresponds to how well Λ projects onto the base space M. At a point where the tangent space of Λ projects surjectively onto M, the L^∞ norm of u is $\mathcal{O}(1)$. The only other possibility that could arise in the regular case is a turning point; that is, a point where the function r has a nondegenerate critical point when restricted to a geodesic (for the sphere, these are the points of maximal and minimal latitude on a given great circle). The behavior of u near the turning points is similar to that of the Airy function, and its L^∞ norm is of order $h^{-1/6}$ by a variant of Van der Corput's lemma.

If one is unfamiliar with Lagrangian distributions, the simple model case to consider would be the eigenfunctions of the Laplacian $h^2 D_x^2$ on the circle $\mathbb{S}^1 = \mathbb{R}/(2\pi\mathbb{Z})$. Those are given by $e^{i\lambda x/h}$, with the eigenvalue λ^2. (Note that only a discrete set of λ is possible here – this is a baby version of the quantization condition mentioned in the next section.) The corresponding symbol is $p(x, \xi) = \xi^2$ and the corresponding Lagrangian would be $\{x \in \mathbb{S}^1, \ \xi = \lambda\}$; it projects surjectively onto the x variables, which corresponds to the fact that we do not need any integration variables θ in the formula (4.1) to define eigenfunctions, and to the fact that the L^∞ norm of eigenfunctions is bounded by a constant.

For $\alpha = 0$, a new kind of problem arises – the intersection of Λ with the fiber of T^*M at a pole is a not a point, but a circle consisting of all unit covectors at the pole, leading to a loss of $h^{-1/2}$ in the L^∞ norm. This problem disappears and we get back the $h^{1/6}$ estimate if either α is away from zero or we are away from the pole, which is reflected in (3.10). If $\alpha = 0$, the turning points are located at the poles; therefore, away from the poles we get an $\mathcal{O}(1)$ estimate. The blow-up rate of this estimate as we approach a pole is quantified by (3.11).

In case (2) we can separate out the φ variable (as Λ does not pass through any poles) and obtain a one-dimensional problem; then u is a low-lying eigenfunction of a Schrödinger operator with a potential well. The bottom of the well eigenfunctions (those with $g(0) - \alpha^2 = \mathcal{O}(h)$) are approximated by the Gaussian $h^{-1/4}e^{-r^2/(2h)}$; we see that they are

$$\mathcal{O}(\min(h^{-1/4}, r^{-1/2})).$$

This explains the first part of (3.7) and (3.8). However, if $g(0) - \alpha^2$ is bounded away from zero, we are away from the equator and thus back to case (1), which explains the $h^{-1/6}$ term in (3.7).

4. Preliminaries

4.1. Semiclassical Lagrangian distributions.
In this subsection, we briefly review the local theory of semiclassical Lagrangian distributions; see for example [2],[15, Chapter 6],

[16, Chapter 8] or [29, Section 2.3] for a detailed account, and [19, Section 25.1] or [14, Chapter 11] for the presentation in the closely related microlocal case.

Assume that X is a manifold and $\Phi(x, \theta)$ is a smooth real-valued function defined on an open set $U_\Phi \subset X \times \mathbb{R}^m$. Define the critical set $C_\Phi \subset U_\Phi$ by

$$(x, \theta) \in C_\Phi \iff \partial_\theta \Phi(x, \theta) = 0.$$

The function Φ is called a (nondegenerate) phase function, if for each $(x, \theta) \in C_\Phi$, the differentials $d(\partial_{\theta_1} \Phi), \ldots, d(\partial_{\theta_m} \Phi)$ are linearly independent at (x, θ). If this is the case, the set

$$\Lambda_\Phi = \{(x, \partial_x \Phi(x, \theta)) \mid (x, \theta) \in C_\Phi\} \subset T^*X$$

is an (immersed) Lagrangian submanifold. We say that Φ generates Λ_Φ; in general, if $\Lambda \subset T^*X$ is a Lagrangian submanifold and $(x, \xi) \in \Lambda_\Phi \subset \Lambda$, then we say that Φ generates Λ near (x, ξ). For each Lagrangian submanifold Λ and each $(x, \xi) \in \Lambda$, there exists a phase function generating Λ near (x, ξ); however, such phase function is not unique.

If $\Lambda \subset T^*X$ is an (embedded) Lagrangian submanifold, Φ is a phase function generating Λ near some point, and $a(x, \theta; h) \in \mathcal{C}_c^\infty(U_\Phi)$ is bounded in $\mathcal{C}_{x,\theta}^\infty$ uniformly in h, we can define the h-dependent family of smooth functions

$$u(x; h) = h^{-m/2} \int e^{i\Phi(x,\theta)/h} a(x, \theta) \, d\theta. \tag{4.1}$$

Here the factor $h^{-m/2}$ is chosen so that $\|u(x; h)\|_{L^2}$ is bounded by a certain $\mathcal{C}^\infty$ seminorm of a and equivalent to the L^2 norm of $a|_{C_\Phi}$, modulo $\mathcal{O}(h)$ terms. We call $u(x; h)$ a (semiclassical compactly microlocalized) Lagrangian distribution associated to Λ. This family is microlocalized on Λ_Φ in the following sense:

$$b \in \mathcal{C}_c^\infty(T^*X), \ b|_{\mathrm{neigh}(\Lambda_\Phi)} = 0 \implies \|b^w(x, hD_x)u(x; h)\|_{L^2} = \mathcal{O}(h^\infty).$$

More generally, we call $u(x; h)$ a Lagrangian distribution associated to Λ, if it is the sum of finitely many expressions of the form (4.1), for different phase functions parametrizing Λ, and an $\mathcal{O}_{\mathcal{C}_c^\infty}(h^\infty)$ remainder.

If $u(x; h)$ is a Lagrangian distribution associated to Λ, Φ is a phase function generating Λ, and u is microlocalized in a compact subset of Λ_Φ, then we can always write u in the form (4.1) for some amplitude a modulo an $\mathcal{O}(h^\infty)$ remainder. (The general case can always be reduced to this one by a microlocal partition of unity, if we know enough phase functions to cover the whole Λ.) In other words, if two phase functions locally generate the same Lagrangian, then oscillatory integrals (4.1) associated to one phase function can be written in the form (4.1) using the other phase function as well. However, the formulas relating even the principal parts of the amplitudes corresponding to different phase functions are quite complicated; obtaining a geometric interpretation for these formulas is the subject of global theory of Lagrangian distributions. This global theory is needed to produce quantization conditions that we use below; however, as we are only interested in

estimating the resulting eigenfunctions and in some rough properties of the spectrum given by quantization conditions, we do not use the global theory directly.

4.2. **Specific generating functions and L^∞ estimates.** In this subsection, we assume that $\Lambda \subset T^*\mathbb{R}^n$ is a Lagrangian submanifold and $(x^0, \xi^0) \in \Lambda$.

Proposition 4.1. *Take $0 \leq m \leq n$ and denote $x = (x', x''), \xi = (\xi', \xi'')$, where $x', \xi' \in \mathbb{R}^m$ and $x'', \xi'' \in \mathbb{R}^{n-m}$. Assume that $T_{(x^0, \xi^0)}\Lambda$ projects surjectively onto the (x', ξ'') variables. Then there exists a function $S(x', \xi'')$ such that near (x^0, ξ^0), Λ is given by*

$$\{\xi' = -\partial_{x'}S(x', \xi''), \ x'' = \partial_{\xi''}S(x', \xi'')\}.$$

Consequently, Λ is generated near (x^0, ξ^0) by the phase function

$$\Phi(x, \theta) = x'' \cdot \theta - S(x', \theta), \ \theta \in \mathbb{R}^{n-m}.$$

Proof. For the reader's convenience we recall the well known argument. We can write Λ locally as a graph $\{\xi' = -F(x', \xi''), \ x'' = G(x', \xi'')\}$. The restriction of the symplectic form $d\xi \wedge dx$ to Λ is zero; therefore, the restriction of the 1-form $\alpha = x'' \, d\xi'' - \xi' \, dx'$ to Λ is closed. Therefore, there exists a function $S(x', \xi'')$ such that $\alpha = dS$ when restricted to Λ. However, if we use (x', ξ'') as a coordinate system on Λ, then $\alpha = G \, d\xi'' + F \, dx'$ and $dS = \partial_{\xi''}S \, d\xi'' + \partial_{x'}S \, dx'$; it follows that $F = \partial_{x'}S$ and $G = \partial_{\xi''}S$. $\square$

In one special case Λ can be locally parametrized by x; the corresponding Lagrangian distributions satisfy the best L^∞ estimate possible:

Proposition 4.2. *Assume that the tangent space to Λ at each point projects surjectively onto the x variables. Then, each Lagrangian distribution $u(x; h)$ associated to Λ satisfies $\|u(x; h)\|_{L^\infty} = \mathcal{O}(1)$.*

Proof. Putting $x' = x$ and $x'' = \emptyset$ in Proposition 4.1, we get a phase function $\Phi(x) = -S(x)$. The expression (4.1) for this phase function has the form $e^{-iS(x)/h}a(x; h)$ and is trivially $\mathcal{O}(1)$ in L^∞. $\square$

If Λ does not satisfy the condition of Proposition 4.2, then the associated Lagrangian distributions can have L^∞ norm as large as $h^{-n/2}$. However, in some cases we are still able to find a phase function satisfying some additional conditions that will ensure a better bound. We will in particular use the following estimate for the case when θ is one-dimensional:

Proposition 4.3. *Assume that $\Phi(x, \theta), (x, \theta) \in U \subset \mathbb{R}^n \times \mathbb{R}$, is a phase function and*

$$|\partial_\theta \Phi(x, \theta)| + |\partial_\theta^2 \Phi(x, \theta)| + |\partial_\theta^3 \Phi(x, \theta)| > 0 \text{ for each } (x, \theta) \in U.$$

Then each oscillatory integral $u(x; h)$ of the form (4.1) satisfies

$$\|u\|_{L^\infty} = \mathcal{O}(h^{-1/6}).$$

Proof. This follows from the van der Corput's Lemma as presented in [30, Proposition VIII.2]. We can also argue using more precise asymptotics which are relevant when precise information at the caustic is needed: If the amplitude $a(x, \theta; h)$ is supported in a region where $|\partial_\theta \Phi| + |\partial_\theta^2 \Phi| > 0$, then we can apply the stationary phase method to (4.1) to get $\|u\|_{L^\infty} = \mathcal{O}(1)$. (Alternatively, the corresponding piece of the Lagrangian Λ_Φ will satisfy the condition of Proposition 4.2.) If at (x_0, θ_0) we have $\partial_\theta \Phi(x_0, \theta_0) = \partial_\theta^2 \Phi(x_0, \theta_0) = 0$ and $\partial_\theta^3 \Phi(x_0, \theta_0) \neq 0$, then we can apply [18, Theorem 7.7.18] which shows that for a supported near (x_0, θ_0), there exist smooth functions a_1, a_2, b such that a_1, b are real-valued and

$$h^{-\frac{1}{2}} \int e^{i\Phi(x,\theta)/h} a(x, \theta) \, d\theta = h^{-1/6} e^{ib(x)/h} Ai(h^{-\frac{2}{3}} a_1(x)) a_2(x) + \mathcal{O}(h^{\frac{1}{6}}).$$

Since the Airy function Ai is bounded, this expression is $\mathcal{O}(h^{-1/6})$. A partition of unity argument finishes the proof of the proposition. $\square$

4.3. **L^∞ estimates for potential wells.** Assume that $V(x)$, $x \in \mathbb{R}$, is a smooth real-valued potential such that:

- for each N, there exists a constant C_N such that $|\partial_x^N V(x)| \leq C_N(1 + |x|^2)$ for all x;
- there exists a constant $\varepsilon_V > 0$ such that $V(x) \geq \varepsilon_V |x|^2$ for all x;
- $V(0) = V'(0) = 0$, $V''(0) = 2c^2$ for some $c > 0$, and $\pm V'(x) > 0$ for $\pm x > 0$.

Consider the semiclassical Schrödinger operator

$$P(h) = h^2 D_x^2 + V(x);$$

let $p(x, \xi) = \xi^2 + V(x)$ be the corresponding principal symbol. Since $p(x, \xi) \to \infty$ as $|x| + |\xi| \to \infty$, it follows [31, Section 6.3] (see also [27, Proposition III.5]) that $P(h)$ is a self-adjoint operator on $L^2(\mathbb{R})$ with compact resolvent and therefore posesses a complete orthonormal system of eigenfunctions.

For $\lambda > 0$, p does not have critical points on $p^{-1}(\lambda)$ and thus produces a completely integrable one-dimensional Hamiltonian system. The following is a corollary of the spectral theory of self-adjoint operators corresponding to quantum completely integrable systems (see [15, Section 2.7] or [29, Theorème 5.1.11]):

Proposition 4.4. *Let K be a compact subset of $(0, \infty)$. Then for h small enough, the eigenvalues λ of $P(h)$ in K are simple and given by the quantization condition*

$$S(\lambda) = (2j + 1)\pi h + \mathcal{O}(h^2), \; j \in \mathbb{Z}, \tag{4.2}$$

where $S(\lambda)$ is the action functional:

$$S(\lambda) = \oint_{p(x,\xi)=\lambda} \xi \, dx$$

and the curve $p^{-1}(\lambda)$ is oriented clockwise. In particular, S is smooth and $S'(\lambda) > 0$ everywhere. Moreover, each L^2 normalized joint eigenfunction corresponding to an eigenvalue
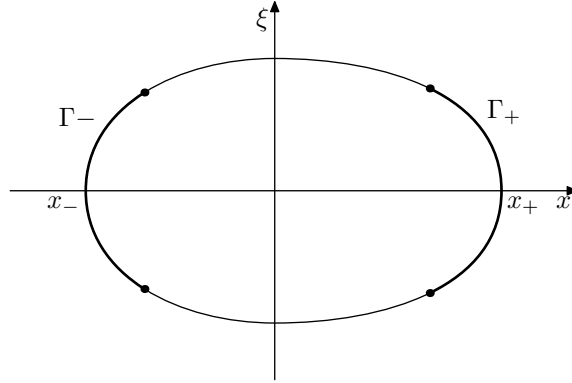
FIGURE 4.1. The curve $p^{-1}(\lambda)$ and the regions $\Gamma_\pm$

$\lambda \in K$ *is a Lagrangian distribution associated to* $p^{-1}(\lambda)$ *(that is, it admits a parametrization by integrals of the form* (4.1) *with estimates on the symbol and the remainder uniform in* λ).

Remark. The quantization condition (4.2) is valid for the eigenvalues close to zero as well, if we add the requirement that $j \geq 0$. This can be proved by conjugating $P(h)$ by a semiclassical Fourier integral operator to a function of the harmonic oscillator; it follows from the analysis in [6] and can be found for example in [29, Théorème 5.2.4].

For λ away from zero, we get the following estimate on eigenfunctions of $P(h)$:

Proposition 4.5. *Let* K *be a compact subset of* $(0, \infty)$. *Then each* L^2*-normalized eigenfunction* v_λ *with eigenvalue* $\lambda \in K$ *satisfies*

$$\|v_\lambda\|_{L^\infty} = \mathcal{O}(h^{-1/6}).$$

Proof. By Proposition 4.4, v_λ is a Lagrangian distribution associated to the curve $p^{-1}(\lambda)$. Let $\Gamma_\pm$ be the intersections of $p^{-1}(\lambda)$ with small neighborhoods of the turning points $(x_\pm(\lambda), 0)$; here $x_-(\lambda) < x_+(\lambda)$ are the two roots of the equation $V(x) = \lambda$. Away from $\Gamma_\pm$, the curve $p^{-1}(\lambda)$ projects diffeomorphically onto the x axis; therefore, by Proposition 4.2, Lagrangian distributions associated to $p^{-1}(\lambda) \setminus (\Gamma_+ \cup \Gamma_-)$ have L^∞ norm $\mathcal{O}(1)$. It remains to consider the case of a Lagrangian distribution associated to, say, Γ_+. However, Γ_+ can be parametrized by ξ and thus by Proposition 4.1 it has a generating function of the form $\Phi(x, \xi) = x\xi - T(\xi)$. We can calculate that $\partial_\xi^3 T(\xi)$ is nonvanishing; it remains to apply Proposition 4.3. $\qquad\square$

Using Proposition 4.5 together with elliptic estimates, we get L^∞ bounds on functions satisfying the analogue of (3.2):

Proposition 4.6. *Assume $\lambda \in \mathbb{R}$ lies in a fixed compact set K and $u \in \mathcal{C}_c^\infty(\mathbb{R})$ satisfies*

$$\|u\|_{L^2} \leq C_0, \quad \|(P(h) - \lambda)u\|_{L^2} \leq C_0 h, \tag{4.3}$$

for some constant C_0. Then:

(1) *If $\chi \in \mathcal{C}_c^\infty(T^*\mathbb{R})$ is equal to 1 near $p^{-1}(\lambda)$, then*

$$\|(1 - \chi^w(x, hD_x))u\|_{L^\infty} = \mathcal{O}_{\chi,\lambda}(1).$$

In particular, if $\pi : T^*\mathbb{R} \to \mathbb{R}$ is the projection map, then

$$x \notin \pi(p^{-1}(\lambda)) \implies |u(x)| = \mathcal{O}_{x,\lambda}(1). \tag{4.4}$$

The constants in $\mathcal{O}(\cdot)$ are locally uniform in x, λ.

(2) *If $|\lambda| \geq \delta$ for some constant $\delta > 0$, then*

$$\|u\|_{L^\infty} = \mathcal{O}_\delta(h^{-1/6}).$$

(3) *If $|x| \geq \delta$ for some constant $\delta > 0$, then*

$$|u(x)| = \mathcal{O}_\delta(h^{-1/6}).$$

Remark. We do not assume the localization condition (3.4) and work only under the hypothesis (4.3). That is because we will apply the proposition to a rescaled version of the original u which no longer satisfies (3.4).

Proof. 1. The operator $P(h) - \lambda$ is elliptic with respect to the order function $m(x, \xi) = 1 + x^2 + \xi^2$ on $\operatorname{supp}(1 - \chi)$; therefore, the elliptic parametrix construction (as in [31, Section 4.5] or [8, Proposition 8.6], with minor adjustments found for example in [12, Proposition 5.1]) gives a symbol $q \in S(m^{-1})$ such that

$$q^w(x, hD_x)(P(h) - \lambda) = 1 - \chi^w(x, hD_x) + \mathcal{O}_{L^2 \to L^2}(h^\infty).$$

Therefore, by Sobolev embedding

$$\begin{aligned}
\|(1 - \chi^w(x, hD_x))u\|_{L^\infty} &\lesssim \|(1 - \chi^w(x, hD_x))u\|_{H^1} \lesssim h^{-1}\|(1 - \chi^w(x, hD_x))u\|_{H_h^1} \\
&\lesssim h^{-1}\|q^w(x, hD_x)(P(h) - \lambda)u\|_{H_h^1} + \mathcal{O}(h^\infty)\|u\|_{L^2} \\
&\lesssim h^{-1}\|(P(h) - \lambda)u\|_{L^2} + \mathcal{O}(h^\infty)\|u\|_{L^2} = \mathcal{O}(1).
\end{aligned}$$

To show (4.4), it suffices to apply the previous inequality with χ vanishing near $\pi^{-1}(x)$, but equal to 1 near $p^{-1}(\lambda)$.

2. If $\lambda \leq -\delta$, then $p^{-1}(\lambda) = \emptyset$ and thus part 1 of this proposition applies with $\chi \equiv 0$. We now assume that $\lambda \geq \delta$. Take $\tilde{\chi} \in \mathcal{C}_c^\infty(\lambda/2, 2\lambda)$ equal to 1 near λ; then the operator $\tilde{\chi}(P(h))$, defined by means of functional calculus, is pseudodifferential, compactly microlocalized, and equal to the identity microlocally near $p^{-1}(\lambda)$ [8, Chapter 8]. By part 1 of this proposition, we have

$$u = \tilde{\chi}(P(h))u + \mathcal{O}_{L^\infty}(1);$$

it remains to estimate $\|\tilde{\chi}(P(h))u\|_{L^\infty}$. Let $v_1, v_2, \ldots \in \mathcal{C}^\infty$ be the orthonormal basis of $L^2(\mathbb{R})$ consisting of eigenfunctions of $P(h)$ with eigenvalues $\lambda_1, \lambda_2, \ldots$. If

$$u = \sum_j c_j v_j,$$

then by (4.3),

$$\sum_j (|\lambda_j - \lambda| + h)^2 |c_j|^2 \leq 4C_0^2 h^2;$$

$$\|\tilde{\chi}(P(h))u\|_{L^\infty} \leq \sum_{\lambda/2 \leq \lambda_j \leq 2\lambda} |c_j| \cdot \|v_j\|_{L^\infty}$$

$$\leq 2C_0 h \left(\sup_{\lambda/2 \leq \lambda_j \leq 2\lambda} \|v_j\|_{L^\infty} \right) \cdot \left(\sum_{\lambda/2 \leq \lambda_j \leq 2\lambda} (|\lambda_j - \lambda| + h)^{-2} \right)^{1/2}.$$

The sum in the second factor is $\mathcal{O}(h^{-1})$ by the quantization condition (4.2). The supremum in the first factor is $\mathcal{O}(h^{-1/6})$ by Proposition 4.5.

3. There exists $\delta_1 > 0$, depending on δ, such that if $|x| \geq \delta$ and $|\lambda| \leq \delta_1$, then $x \notin \pi(p^{-1}(\lambda))$. The case $|\lambda| \leq \delta_1$ is then handled by (4.4), while the case $|\lambda| \geq \delta_1$ is handled by part 2 of this proposition. $\qquad\square$

5. Proof of Theorem 2

5.1. **Regular case.** We have

$$g(r)[-h^2\Delta - 1] = h^2[D_r^2 + D_\varphi^2] - g(r).$$

We then separate out the φ variable: if $u(r, \varphi) = u(r)e^{i\alpha\varphi/h}$, then

$$u(r) \in \mathcal{C}_c^\infty(-\varepsilon, \varepsilon), \quad \|u\|_{L^2} = \mathcal{O}(1), \quad \|(h^2 D_r^2 + \alpha^2 - g(r))u\|_{L^2} = \mathcal{O}(h). \qquad (5.1)$$

Now, for ε small enough,

$$g(r) = V_0 - V(r + x_0), \quad |r| < \varepsilon$$

where a potential $V(x)$ satisfies conditions of Section 4.3. Moreover, $|x_0| > 2\varepsilon$. If $\tilde{u}(x) = u(x - x_0)$, then $\|(P(h) - V_0 + \alpha^2)\tilde{u}\|_{L^2} = \mathcal{O}(h)$, where

$$P(h) = h^2 D_x^2 + V(x).$$

However, $\tilde{u}$ is supported away from zero; therefore, by part 3 of Proposition 4.6, we find

$$\|u\|_{L^\infty} = \mathcal{O}(h^{-1/6})$$

as required.

5.2. **Case of an equator.** As in the regular case, we can separate out the φ variable and obtain (5.1). Now, take a potential $V(r)$ satisfying conditions of Section 4.3 such that $V(r) = g(0) - g(r)$ for $|r| < \varepsilon$. Then, $\|(P(h) - \lambda)u\|_{L^2} = \mathcal{O}(h)$, with

$$P(h) = h^2 D_r^2 + V(r), \quad \lambda = g(0) - \alpha^2.$$

For λ or r bounded away from zero, we can argue similarly to the regular case. Indeed, (3.7) follows from parts 2 and 3 of Proposition 4.6. To obtain the estimate (3.8) for r bounded away from zero, we use (4.4): the condition in (3.8) means that $|\lambda| \leq \varepsilon r^2$ so that $r \notin \pi(p^{-1}(\lambda))$.

If both λ and r are close to zero, we will use the natural rescaling of the quantum harmonic oscillator: for some constant C_1 and $C_1^{-1}h \leq \beta \leq C_1$, define

$$(T_\beta f)(\tilde{r}) = \beta^{1/4} f(\beta^{1/2}\tilde{r}), \ f \in L^2(\mathbb{R});$$

$$h_\beta = \beta^{-1}h, \ \lambda_\beta = \beta^{-1}\lambda,$$

$$V_\beta(\tilde{r}) = \beta^{-1}V(\beta^{1/2}\tilde{r}), \quad P_\beta(h_\beta) = h_\beta^2 D_{\tilde{r}}^2 + V_\beta(\tilde{r}).$$

We note that if $V(r) = r^2$, then

$$P_\beta(h_\beta) = h_\beta^2 D_{\tilde{r}}^2 + V(\tilde{r});$$

that is, the operator does not change. For a general potential we get closer to the harmonic oscillator as $\beta \to 0$. Moreover, the potential V_β satisfies conditions of Section 4.3 uniformly in β. Indeed, the only nontrivial part is verifying the first of these conditions for $N = 0, 1$, and this follows from the fact that $V(r) = \mathcal{O}(r^2)$ and $V'(r) = \mathcal{O}(|r| + r^2)$. Therefore, the constants in the estimates of Proposition 4.6 do not depend on β.

The operator T_β is unitary on $L^2(\mathbb{R})$ and

$$P_\beta(h_\beta) - \lambda_\beta = \beta^{-1}T_\beta(P(h) - \lambda)T_\beta^*,$$

Therefore, if $u_\beta = T_\beta u$, then

$$\|u_\beta\|_{L^2} = \mathcal{O}(1), \ \|(P_\beta(h_\beta) - \lambda_\beta)u_\beta\|_{L^2} = \mathcal{O}(h_\beta).$$

For $\varepsilon > 0$ small enough, we have $P(h) \geq \varepsilon(h^2 D_r^2 + r^2)$; the ground state of the quantum harmonic oscillator, which by scaling invariance is known to be bound from below by εh ([31, Section 6.1]), giving thus $P(h) \geq \varepsilon h$. Therefore, we can assume that $|\lambda| \geq \varepsilon h$ – indeed, if $|\lambda| \leq \varepsilon h$, then we can replace λ by εh and obtain same conditions on u and stronger conclusions (except for (3.8) in case when $r^2 < h$ – however, in this case $r^{-1/2} \geq h^{-1/4}$ and (3.8) follows from (3.7)). Now, consider the following two subcases:

(1) $|\lambda| < \varepsilon r^2$. Since $|\lambda| \geq \varepsilon h$, we have $|r| > h^{1/2}$. Put $\beta = r^2$; then $u(r) = r^{-1/2}u_\beta(\pm 1)$ (depending on the sign or r) and $|\lambda_\beta| < \varepsilon$. For ε small enough, $\tilde{r} = \pm 1$ does not lie in the projection of $\{\xi^2 + V_\beta(\tilde{r}) = \lambda_\beta\}$; therefore, (4.4) applies, giving $|u_\beta(\pm 1)| = \mathcal{O}(1)$ and thus $|u(r)| = \mathcal{O}(r^{-1/2})$; this proves (3.8), which implies (3.7) in the present case.

(2) $|\lambda| \geq \varepsilon r^2$, where ε is chosen as in case (1). Put $\beta = |\lambda|$; then $\lambda_\beta = \pm 1$ depending on the sign of λ and $u(r) = |\lambda|^{-1/4} u_\beta(\tilde{r})$, where $\tilde{r} = |\lambda|^{-1/2} r$ is bounded. Now, part 2 of Proposition 4.6 applies for $\lambda_\beta = 1$ and part 1 of the same proposition applies for $\lambda_\beta = -1$ (with $\chi = 0$), giving $|u_\beta(\tilde{r})| = \mathcal{O}(h_\beta^{-1/6})$ and thus $|u(r)| = \mathcal{O}(h^{-1/6}|\lambda|^{-1/12})$. This proves (3.7) as the third term in the minimum expression is controlled by each of the other two in the present case.

5.3. **Case of a pole.** Define the operators $P(h) = -h^2 \Delta$ and $Q(h) = hD_\varphi$; let p, q be the corresponding semiclassical principal symbols; for $\lambda \geq 0$, consider the compact set

$$\Lambda(\lambda, \alpha) = \{p = \lambda, \ q = \alpha\} \subset T^*M.$$

We start by an analogue of part 1 of Proposition 4.3:

Proposition 5.1. *Assume that $X(h)$ is a compactly microlocalized pseudodifferential operator on $\mathbb{S}^2$ microlocally equal to the identity near $\Lambda(1, \alpha)$. Then*

$$\|(1 - X(h))u\|_{L^\infty} = \mathcal{O}(1).$$

Proof. Using the localization assumption (3.4), a microlocal partition of unity, and the elliptic parametrix construction (see the proof of Proposition 4.6) we can write

$$(1 - X(h))u = A(h)(P(h) - 1)u + B(h)(Q(h) - \alpha)u + \mathcal{O}_{\mathcal{C}^\infty}(h^\infty), \qquad (5.2)$$

where both $A(h)$ and $B(h)$ are pseudodifferential and compactly microlocalized, and thus act $L^2 \to L^\infty$ with norm $\mathcal{O}(h^{-1})$ [31, Theorem 7.10]. We now use (3.2) and (3.3) in (5.2) and that proves the proposition. $\qquad \square$

If (ξ, η) are the momenta corresponding to the coordinates $(x, y) = (r \cos \varphi, r \sin \varphi)$, then

$$p = g(x^2 + y^2)^{-1}(\xi^2 + \eta^2), \ q = x\eta - y\xi.$$

Therefore, if α is bounded away from zero and ε is small, then the projection of $\Lambda(1, \alpha)$ onto M does not intersect $U_\varepsilon = \{r < \varepsilon\}$ and $\|u\|_{L^\infty} = \mathcal{O}(1)$ by Proposition 5.1. Therefore, we may assume that $|\alpha| < \varepsilon$. The symbols p and q have linearly independent differentials on $\Lambda(1, 0)$, and they Poisson commute; therefore, $\Lambda(\lambda, \alpha)$ is a Lagrangian torus for small α and λ close to 1. The joint spectrum of $P(h), Q(h)$ near $(1, 0)$ obeys a quantization condition; in particular, the spectrum of $P(h)$ restricted to the eigenspace $\{Q(h) = \alpha\}$ is approximated by a formula similar to (4.2). (See [5] or [29, Theorème 5.1.11].) Similarly to the proof of part 2 of Proposition 4.6, we reduce the problem to the following

Proposition 5.2. *If $\varepsilon > 0$ is small enough, then each L^2 normalized joint eigenfunction v for $P(h), Q(h)$ with eigenvalue (λ, α), where $|\lambda - 1|, |\alpha| < \varepsilon$, satisfies the following estimates for $|r| < \varepsilon$:*

$$|v(r, \varphi)| = \mathcal{O}(\min(h^{-1/2}, h^{-1/6} r^{-1/3}, h^{-1/6}|\alpha|^{-1/3})); \qquad (5.3)$$

$$|v(r, \varphi)| = \mathcal{O}(r^{-1/2}) \ \textit{for } |\alpha| < \varepsilon r. \qquad (5.4)$$

Proof. As follows from the proof of the quantization condition, $v(r, \varphi)$ is a Lagrangian distribution associated to $\Lambda(\lambda, \alpha)$. Using a microlocal partition of unity and the rotational symmetry of the problem, we can reduce to the case when v is microlocalized in $\{\xi > \eta/2, \; |r| < \varepsilon\}$. However, in this region $\Lambda(\lambda, \alpha)$ can be parametrized by x, η, since the matrix of derivatives of p, q in y, ξ is nondegenerate for $x = y = 0$; by Proposition 4.1, $\Lambda(\lambda, \alpha)$ is generated by a phase function of the form $\Phi = y\eta - S(x, \eta; \lambda, \alpha)$, where S is a smooth function. We recall that this means that on $p = \lambda$ and $q = \alpha$, $\xi = -\partial_x S$ and $y = \partial_\eta S$.

From this we calculate

$$\partial_x S(0, \eta, \lambda, \alpha) = -\sqrt{g(y^2)\lambda - \eta^2} = -\sqrt{\lambda g(0) - \eta^2} + \mathcal{O}(\alpha^2) \,,$$

$$\partial_\eta S(0, \eta, \lambda, \alpha) = -\alpha/\xi = -\frac{\alpha}{\sqrt{\lambda g(0) - \eta^2}} + \mathcal{O}(\alpha^2);$$

if we normalize S by the condition $S(0, 0, \lambda, \alpha) = 0$, then $S(0, \eta, \lambda, 0) = 0$, and

$$\partial_\alpha S(0, \eta, \lambda, 0) = \partial_\alpha S(0, 0, \lambda, 0) + \int_0^\eta \partial_\alpha \partial_\eta S(0, t, \lambda, 0) dt = -\int_0^\eta \frac{dt}{\sqrt{\lambda g(0) - t^2}}$$

$$= -\arcsin(\eta/\sqrt{\lambda g(0)}) \,.$$

Now, put $(r, \alpha) = (s\tilde{r}, s\tilde{\alpha})$ with $\tilde{r}^2 + \tilde{\alpha}^2 = 1$ and define $\tilde{h} = h/s$; we get by (4.1),

$$v(r, \varphi; \lambda, \alpha, h) = h^{-1/2} \int \exp\left(i(s\tilde{r}\eta \sin\varphi - S(s\tilde{r}\cos\varphi, \eta; \lambda, s\tilde{\alpha}))/h\right) a \, d\eta$$

$$= s^{-1/2}\tilde{h}^{-1/2} \int \exp\left(i\tilde{\Phi}(\eta, s, \tilde{r}, \tilde{\alpha}, \varphi, \lambda)/\tilde{h}\right) a \, d\eta,$$

where

$$\tilde{\Phi} = \tilde{r}(\eta \sin\varphi + \sqrt{\lambda g(0) - \eta^2}\cos\varphi) + \tilde{\alpha}\arcsin(\eta/\sqrt{\lambda g(0)}) + \mathcal{O}(s) \,,$$

and

$$a = a(s\tilde{r}\cos\varphi, s\tilde{r}\sin\varphi, \eta) \,, \quad a \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^3) \,.$$

For s small enough, $\tilde{\Phi}$ satisfies the condition of Proposition 4.3. Indeed, under the change of variables $\eta = \sqrt{\lambda g(0)}\sin\theta$ we get $\tilde{\Phi} = \tilde{r}\sqrt{\lambda g(0)}\cos(\theta - \varphi) + \tilde{\alpha}\theta + \mathcal{O}(s)$. Therefore,

$$|v| = \mathcal{O}(s^{-1/2}\tilde{h}^{-1/6}) = \mathcal{O}(s^{-1/3}h^{-1/6}).$$

This proves (5.3) (note that the bound $\mathcal{O}(h^{-1/2})$ follows directly from the integral representation). To show (5.4), note that for $|\alpha| < \varepsilon r$, $\tilde{\alpha}$ is close to zero; then $\tilde{\Phi}$ is a Morse function and we can use stationary phase in place of Proposition 4.3 (see also the beginning of the proof of that proposition). $\qquad\square$

Université Paris-Nord for its generous hospitality in the Spring 2011 when this paper was written.

References

[1] P. Abrial, Y. Moudden, J. Starck, J. Fadili, J. Delabrouille, and M. Nguyen, *C.M.B. data analysis and sparsity,* Stat. Methodol. **5**(2008), 289–298.

[2] I. Alexandrova, *Semi-Classical wavefront set and Fourier integral operators.* Can. J. Math. **60**(2008), 241–263.

[3] N. Burq, P. Gérard, and N. Tzvetkov, *Restrictions of the Laplace–Beltrami eigenfunctions to submanifolds,* Duke. Math. J. **138**(2007), 445–486.

[4] E. Candes, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,* IEEE Trans. Inform. Theory **52**(2006), 489–509.

[5] A.-M. Charbonnel, *Spectre conjoint d'opérateurs pseudodifférentiels qui commutent,* Ann. Fac. Sci. Toulouse Math. **5**(1983), 109–147.

[6] Y. Colin de Verdière, *Spectre conjoint d'opérateurs pseudo-différentiels qui commutent. II. Le cas intégrable,* Math. Z. **171**(1980), 51–73.

[7] A. Cohen, W. Dahmen, and R. DeVore *Compressed sensing and best k-term approximation,* J. Amer. Math. Soc. **22** (2009), no. 1, 211–231

[8] M. Dimassi and J. Sjöstrand, *Specral asymptotics in the semi-classical limit,* Cambridge University Press, 1999.

[9] D. Donoho, *Compressed sensing,* IEEE Trans. Inform. Theory **52**(2006), 1289–1306.

[10] D. Donoho and P. Starck, *Uncertainty principles and signal recovery,* SIAM J. Appl. Math **49** 3 (1989), 906–931.

[11] D. Donoho and X. Huo, *Uncertainty principles and ideal atomic decompositions,* IEEE Trans. Inform. Theory **47**(2001), 2845–2862.

[12] S. Dyatlov, *Quasi-normal modes and exponential energy decay for the Kerr–de Sitter black hole,* Comm. Math. Phys. **306**(2011), 119–163.

[13] S. Foucart, A. Pajor and H. Rauhut, *The Gelfand widths of ℓ_p-balls for $0 < p \le 1$* J. Complexity 26 (2010), no. 6, 629–640

[14] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators: an introduction,* Cambridge University Press, 1994.

[15] V. Guillemin and S. Sternberg, *Geometric asymptotics,* AMS, 1990.

[16] V. Guillemin and S. Sternberg, *Semi-classical analysis,* lecture notes, http://www-math.mit.edu/~vwg/semiclassGuilleminSternberg.pdf

[17] L. Hörmander, *The spectral function of an elliptic operator.* Acta Math. **121**(1968), 193–218.

[18] L. Hörmander, *The Analysis of Linear Partial Differential Operators. I. Distribution Theory and Fourier Analysis,* Springer Verlag, 1983.

[19] L. Hörmander, *The Analysis of Linear Partial Differential Operators. IV. Fourier Integral Operators,* Springer Verlag, 1994.

[20] I. Krasikov, *On the Erdelyi-Magnus-Nevai conjecture for Jacobi polynomias,* Constr. Approx., 2008, 113–125.

[21] A. Martinez *An introduction to semiclassical and microlocal analysis,* Universitext. Springer-Verlag, New York, 2002. viii+190 pp. ISBN: 0-387-95344-2

[22] R. Morris, R. Najmanovich, A. Kahraman, and J. Thornton, *Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparison,* Bioinformatics, **21** 10 (2005), 2347–2355.

[23] H. Rauhut, *Compressive Sensing and Structured Random Matrices,* in Theoretical Foundations and Numerical Methods for Sparse Recovery, Radon Series Comp. Appl. Math, deGruyter, **9**(2010), 1–92.

[24] H. Rauhut, *Stability results for random sampling of sparse trigonometric polynomials,* IEEE Trans. Inform. Theory, **54** 12 (2006). 5661–5670.

[25] H. Rauhut and R. Ward, *Sparse Legendre expansions via ℓ_1-minimization,* Submitted (2010).

[26] H. Rauhut and R. Ward, *Sparse recovery for spherical harmonic expansions,* Proc. SampTA 2011, Singapore, 2011.

[27] D. Robert, *Autour de l'approximation semi-classique,* Progress in Mathematics, 68. Birkhuser Boston, Inc., Boston, MA, 1987. x+329 pp. ISBN: 0-8176-3354-5

[28] M. Rudelson and R. Vershynin, *On sparse reconstruction from Fourier and Gaussian measurements,* Comm. Pure Appl. Math, **61**(2008), 1025–1045.

[29] San Vũ Ngọc, *Systèmes intégrables semi-classiques: du local au global,* Panoramas et Synthèses 22, 2006.

[30] E.M. Stein and T.S. Murphy, *Harmonic Analysis: real variable methods, orthogonality, and oscillatory integrals,* Princeton University Press, 1993.

[31] M. Zworski, *Semiclassical analysis,* Graduate Studies in Mathematics, AMS, to appear.

E-mail address: `nicolas.burq@math.u-psud.fr`

Département de Mathématiques, Université Paris-Sud, 91405 Orsay Cedex, France

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

Department of Mathematics, University of California, Berkeley, CA 94720, USA

E-mail address: `rward@math.utexas.edu`

Mathematics Department, University of Texas at Austin, Austin, TX 78712, USA

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

Department of Mathematics, University of California, Berkeley, CA 94720, USA