Quadrature rules and distribution of points on manifolds

Luca Brandolini Christine Choirat Leonardo Colzani
Giacomo Gigante* Raffaello Seri Giancarlo Travaglini

Abstract

We study the error in quadrature rules on a compact manifold,

$$\left|\sum_{j=1}^{N}\omega_j f\left(z_j\right)-\int_{\mathcal{M}}f(x)dx\right|\leq c\mathcal{D}\left\{z_j\right\}\mathcal{V}\left\{f\right\}.$$

As in the Koksma Hlawka inequality, $\mathcal{D}\left\{z_j\right\}$ is a sort of discrepancy of the sampling points and $\mathcal{V}\left\{f\right\}$ is a generalized variation of the function. In particular, we give sharp quantitative estimates for quadrature rules of functions in Sobolev classes.

Keywords. Quadrature, discrepancy, harmonic analysis

1 Introduction

In what follows, $\mathcal{M}$ is a smooth compact d dimensional Riemannian manifold with Riemannian measure dx, normalized so that the total volume of the manifold is 1, and Δ is the Laplace Beltrami operator. This operator is self-adjoint in $\mathbb{L}^2(\mathcal{M})$, it has a sequence

L. Brandolini: Dipartimento di Ingegneria dell'Informazione e Metodi Matematici, Università di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia; e-mail: luca.brandolini@unibg.it

C. Choirat: Department of Economics, School of Economics and Business Management, Universidad de Navarra, Edificio de Bibliotecas (Entrada Este), 31080 Pamplona (Spain); e-mail: cchoirat@unav.es

L. Colzani: Dipartimento di Matematica e Applicazioni, Edificio U5, Università di Milano Bicocca, Via R.Cozzi 53, 20125 Milano, Italia; e-mail: leonardo.colzani@unimib.it

G. Gigante (corresponding author): Dipartimento di Ingegneria dell'Informazione e Metodi Matematici, Università di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia; e-mail: giacomo.gigante@unibg.it

R. Seri: Università degli Studi dell'Insubria, Dipartimento di Economia, via Monte Generoso 71, 21100 Varese, Italia; e-mail: raffaello.seri@uninsubria.it

G. Travaglini: Dipartimento di Statistica, Edificio U7, Università di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italia; e-mail: giancarlo.travaglini@unimib.it

Mathematics Subject Classification (2010): Primary 41A55; Secondary 11K38, 42C15

of eigenvalues $\{\lambda^2\}$ and an orthonormal complete system of eigenfunctions $\{\varphi_\lambda(x)\}$, $\Delta\varphi_\lambda(x) = \lambda^2\varphi_\lambda(x)$. The eigenvalues, possibly repeated, can be ordered with increasing modulus. In particular, the first eigenvalue is 0 and the associated eigenfunction is 1. An example is the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with the Laplace operator $-\sum \partial^2/\partial x_j^2$, eigenvalues $\left\{4\pi^2 |k|^2\right\}_{k\in\mathbb{Z}^d}$ and eigenfunctions $\{\exp(2\pi ikx)\}_{k\in\mathbb{Z}^d}$. Another example is the sphere $\mathbb{S}^d = \left\{x \in \mathbb{R}^{d+1},\ |x| = 1\right\}$ with normalized surface measure and with the angular component of the Laplacian in the space $\mathbb{R}^{d+1}$, eigenvalues $\{n(n+d-1)\}_{n=0}^{+\infty}$ and eigenfunctions the restriction to the sphere of homogeneous harmonic polynomials in space.

A classical problem is to approximate an integral $\int_{\mathcal{M}} f(x)dx$ with Riemann sums $N^{-1}\sum_{j=1}^{N} f(z_j)$, or weighted analogues $\sum_{j=1}^{N}\omega_j f(z_j)$, and what follows will be concerned with the discrepancy between integrals and sums for functions in Sobolev classes $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with $1 \le p \le +\infty$ and $\alpha > d/p$. The assumption $\alpha > d/p$ guarantees the boundedness and continuity of the function $f(x)$, otherwise $f(z_j)$ may be not defined. As a motivation, assume there exists a decomposition of $\mathcal{M}$ into N disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup ... \cup U_N$ and these pieces have measures N^{-1} and diameters at most $cN^{-1/d}$. Choosing a point z_j in each U_j, one obtains the estimate

$$\left|N^{-1}\sum_{j=1}^{N} f(z_j) - \int_{\mathcal{M}} f(x)dx\right|$$
$$\le \sum_{j=1}^{N}\int_{U_j} |f(z_j) - f(x)|\, dx \le \sup_{|y-x|\le cN^{-1/d}} \{|f(y) - f(x)|\}.$$

In particular, since functions in $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are Hölder continuous of degree $\alpha - d/p$, one obtains

$$\left|N^{-1}\sum_{j=1}^{N} f(z_j) - \int_{\mathcal{M}} f(x)dx\right| \le cN^{-(\alpha-d/p)/d}\, \|f\|_{\mathbb{W}^{\alpha,p}(\mathcal{M})}.$$

On the other hand, it will be shown that suitable choices of the sampling points $\{z_j\}$ improve the exponent $1/p - \alpha/d$ to $-\alpha/d$ and this is best possible. More precisely, the main results in this paper are the following:

(1) For every $d/2 < \alpha < d/2+1$ there exists $c > 0$ such that if $\mathcal{M} = U_1 \cup U_2 \cup ... \cup U_N$ is a decomposition of the manifold in disjoint pieces with measure $|U_j| = \omega_j$, then there exists a distribution of points $\{z_j\}_{j=1}^{N}$ with $z_j \in U_j$ such that for every function $f(x)$ in the Sobolev space $\mathbb{W}^{\alpha,2}(\mathcal{M})$,

$$\left|\sum_{j=1}^{N}\omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx\right| \le c \max_{1\le j\le N}\left\{\text{diameter}(U_j)^{\alpha}\right\} \|f\|_{\mathbb{W}^{\alpha,2}}.$$

(2) Assume that the points $\{z_j\}_{j=1}^{N}$ and the positive weights $\{\omega_j\}_{j=1}^{N}$ give an exact

quadrature for all eigenfunctions with eigenvalues $\lambda^2 < r^2$, that is

$$\sum_{j=1}^{N} \omega_j \varphi_\lambda (z_j) = \int_{\mathcal{M}} \varphi_\lambda(x)dx = \begin{cases} 1 & \textit{if } \lambda = 0, \\ 0 & \text{if } 0 < \lambda < r. \end{cases}$$

Then for every $1 \leq p \leq +\infty$ and $\alpha > d/p$ there exist $c > 0$, which may depend on $\mathcal{M}$, p, α, but is independent of r, $\{z_j\}_{j=1}^N$ and $\{\omega_j\}_{j=1}^N$, such that

$$\left| \sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \leq cr^{-\alpha} \|f\|_{\mathbb{W}^{\alpha,p}} .$$

(3) If $1 \leq p \leq +\infty$ and $\alpha > d/p$, then there exist $c > 0$ and sequences of points $\{z_j\}_{j=1}^N$ and positive weights $\{\omega_j\}_{j=1}^N$ with

$$\left| \sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \leq cN^{-\alpha/d} \|f\|_{\mathbb{W}^{\alpha,p}} .$$

(4) For every $1 \leq p \leq +\infty$ and $\alpha > d/p$ there exists $c > 0$ such that for every distribution of points $\{z_j\}_{j=1}^N$ and numbers $\{\omega_j\}_{j=1}^N$ there exists a function $f(x)$ in $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with

$$\left| \sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \geq cN^{-\alpha/d} \|f\|_{\mathbb{W}^{\alpha,p}} .$$

An explicit example is the following. The torus $\mathbb{T}^d$ can be partitioned into $N = n^d$ congruent cubes with sides $1/n$ and this partition generates the mesh of points $\left(n^{-1}\mathbb{Z}^d\right) \cap \mathbb{T}^d$, which gives an exact quadrature at least for all exponentials $\exp(2\pi ikx)$ with $|k_j| < n$. In this case, (1) and (2) give an upper bound for the error in numerical integration of the order of $N^{-\alpha/d}$. More generally, if a manifold is decomposed into N disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup ... \cup U_N$ with diameters $\leq cN^{-1/d}$, then (1) gives the upper bound $N^{-\alpha/d}$. Moreover, for every $r > 0$ there are approximately cr^d eigenfunctions with eigenvalues $\lambda^2 < r^2$ and one can choose $N \leq cr^d$ nodes $\{z_j\}_{j=1}^N$ and positive weights $\{\omega_j\}_{j=1}^N$ which give an exact quadrature for these eigenfunctions. Then in this case (2) gives the above upper bound $N^{-\alpha/d}$. Hence (1) and (2) imply (3), and by (4) this latter is optimal. When the manifold is a torus or a sphere these results are essentially known, and indeed there is a huge literature on this subject. See [24] for deterministic and stochastic error bounds in numerical analysis. In particular, (3) and (4) for $p = 2$ and for spheres are contained in [7], [15] and [16]. For Besov spaces on spheres a result slightly more precise than (3) is in [17], while a result slightly weaker than (4) for compact two point homogeneous spaces is in [21]. See also [10] and, for a survey on related results, [14] and [19]. Beside the proofs of (1), (2), (3), (4), which are contained in the following section, the paper contains also a final section with a number of further results and remarks. Among them it is proved that if a quadrature rule gives an optimal error in the Sobolev space $\mathbb{W}^{\alpha,2}(\mathcal{M})$, then this quadrature rule is optimal also in all spaces $\mathbb{W}^{\beta,2}(\mathcal{M})$ with $d/2<\beta<\alpha$. Moreover, it is proved that there is a relation between quadrature rules and geometric discrepancy:

(5) If $d\nu(x)$ is a probability measure on $\mathcal{M}$, then the norm of the measure $d\nu(x)-dx$ as a linear functional on $\mathbb{W}^{\alpha,2}(\mathcal{M})$ decreases as α increases. Moreover, if the norm of $d\nu(x)-dx$ on $\mathbb{W}^{\alpha,2}(\mathcal{M})$ is $r^{-\alpha}$,

$$\left|\int_{\mathcal{M}} f(x)d\nu(x)-\int_{\mathcal{M}} f(x)dx\right|\leq r^{-\alpha}\left\|f\right\|_{\mathbb{W}^{\alpha,2}},$$

then for every $d/2<\beta<\alpha$ there exists a constant c which may depend on α, β, $\mathcal{M}$, but is independent of r and $d\nu(x)$, such that

$$\left|\int_{\mathcal{M}} f(x)d\nu(x)-\int_{\mathcal{M}} f(x)dx\right|\leq cr^{-\beta}\left\|f\right\|_{\mathbb{W}^{\beta,2}}.$$

(6) Assume that for some $r\geq 1$ the discrepancy of the probability measure $d\nu(x)$ with respect to the balls $\{B(y,\delta)\}$ with center y and radius δ satisfies the estimates

$$\left|\int_{B(y,\delta)} d\nu(x)-\int_{B(y,\delta)} dx\right|\leq\begin{cases} r^{-d} & \text{if } \delta\leq 1/r,\\ r^{-1}\delta^{d-1} & \text{if } \delta\geq 1/r.\end{cases}$$

Then for every $1\leq p\leq+\infty$ and $\alpha>d/p$, there exists a constant c, which may depend on α and p, but is independent of $d\nu(x)$ and r, such that

$$\left|\int_{\mathcal{M}} f(x)d\nu(x)-\int_{\mathcal{M}} f(x)dx\right|\leq\begin{cases} cr^{-\alpha}\left\|f\right\|_{\mathbb{W}^{\alpha,p}} & \text{if } 0<\alpha<1,\\ cr^{-1}\log(1+r)\left\|f\right\|_{\mathbb{W}^{\alpha,p}} & \text{if } \alpha=1,\\ cr^{-1}\left\|f\right\|_{\mathbb{W}^{\alpha,p}} & \text{if } \alpha>1.\end{cases}$$

Observe that while (1) and (2) hold for specific quadrature rules, (5) is a result for arbitrary quadratures. Actually, (5) is only one way, from α to $\beta<\alpha$. The estimate $r^{-\alpha}$ for an α does not necessarily imply the estimate $cr^{-\beta}$ for $\beta>\alpha$. Moreover, the sets $\{B(y,\delta)\}$ in (6) are not precisely geodesic balls, but level sets of suitable kernels on the manifold. However, for spheres or compact rank one symmetric spaces these sets are geodesic balls. In this case the discrepancy of the measure is the spherical cap discrepancy. See [4] or [23], and for other relations between quadrature and discrepancy on spheres, see also [2]. Finally, we would like to point out that our paper is (almost) self contained, it does not rely on explicit properties of manifolds or special functions, and it may provide a unified vision and simple alternative proofs of some known results.

2 Main results

The eigenfunction expansions of functions and operators are a basic tool in what follows. The system of eigenfunctions $\{\varphi_\lambda(x)\}$ is orthonormal complete in $\mathbb{L}^2(\mathcal{M})$ and to every square integrable function one can associate a Fourier transform and series,

$$\mathcal{F}f(\lambda)=\int_{\mathbb{M}}f(y)\overline{\varphi_\lambda(y)}dy,\quad f(x)=\sum_\lambda \mathcal{F}f(\lambda)\varphi_\lambda(x).$$

Since the Laplace operator is elliptic, the eigenfunctions are smooth and it is possible to extend the definition of Fourier transforms and series to distributions. In particular, the Fourier expansions are always convergent, at least in the topology of distributions. One can write the discrepancy between integral and Riemann sum as a single integral with respect to a measure $d\mu(x)=\sum_{j=1}^{N}\omega_j\delta_{z_j}(x)-dx$, with $\delta_y(x)$ the Dirac measure concentrated at the point y and dx the Riemannian measure,

$$\sum_{j=1}^{N}\omega_j f\left(z_j\right)-\int_{\mathcal{M}}f(x)dx=\int_{\mathcal{M}}f(x)d\mu(x).$$

Then the estimate of the error in the numerical integration reduces to the estimate of the norm of a linear functional $d\mu(x)$ on a space of test functions $f(x)$. Some of the results which follow will be stated for generic finite signed measures $d\mu(x)$, for measures of the form $d\mu(x)=d\nu(x)-dx$ with $d\nu(x)$ a probability measure, and also for atomic probability measures $d\nu(x)=\sum_{j=1}^{N}\omega_j\delta_{z_j}(x)$. The following is an easy and straightforward extension to compact manifolds and p norms of some abstract results for reproducing kernel Hilbert spaces. See e.g. [1], [12], [13].

Theorem 2.1. *Let $\{\psi(\lambda)\}$ be a numeric sequence indexed by the eigenvalues $\{\lambda^2\}$, with $\{\psi(\lambda)\}$ and $\{\psi(\lambda)^{-1}\}$ slowly increasing, that is $|\psi(\lambda)|\leq a\left(1+\lambda^2\right)^{\alpha/2}$ and $|\psi(\lambda)^{-1}|\leq b\left(1+\lambda^2\right)^{\beta/2}$. Let $A(x,y)$ and $B(x,y)$ be distribution kernels with Fourier transforms $\{\psi(\lambda)\}$ and $\{\psi(\lambda)^{-1}\}$,*

$$A(x,y)=\sum_\lambda \psi(\lambda)\varphi_\lambda(x)\overline{\varphi_\lambda(y)},\quad B(x,y)=\sum_\lambda \psi(\lambda)^{-1}\varphi_\lambda(x)\overline{\varphi_\lambda(y)}.$$

Finally, let $f(x)$ be a continuous function and let $d\mu(x)$ be a finite measure on $\mathcal{M}$. Then, if $1\leq p,q\leq+\infty$ and $1/p+1/q=1$,

$$\left|\int_{\mathcal{M}}f(x)d\mu(x)\right|$$
$$\leq\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}A(x,y)f(y)dy\right|^p dx\right\}^{1/p}\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}B(x,y)d\mu(x)\right|^q dy\right\}^{1/q}.$$

In particular, when $p=q=2$ and $B(x,y)=B\left(y,x\right)$ and

$$f(x)=\int_{\mathcal{M}}\int_{\mathcal{M}}B\left(x,y\right)\overline{B\left(y,z\right)}dy\overline{d\mu(z)},$$

then the above inequality reduces to an equality.

Proof. The assumptions $\{\psi(\lambda)\}$ and $\{\psi(\lambda)^{-1}\}$ slowly increasing simply imply that the kernels $A(x,y)$ and $B(x,y)$ are tempered distributions. In what follows the pairing between a test function and a distribution is denoted with an integral, even when the distribution is not a function and the integral is divergent. Let

$$\int_{\mathcal{M}} A(x,y)f(y)dy = \sum_{\lambda} \psi(\lambda)\mathcal{F}f(\lambda)\varphi_\lambda(x),$$
$$\int_{\mathcal{M}} B(x,y)d\mu(y) = \sum_{\lambda} \psi(\lambda)^{-1}\mathcal{F}\mu(\lambda)\varphi_\lambda(x).$$

These operators are one the inverse of the other,

$$\begin{aligned} f(x) &= \int_{\mathcal{M}} B(x,y)\left(\int_{\mathcal{M}} A(y,z)f(z)dz\right)dy \\ &= \int_{\mathcal{M}} A(x,y)\left(\int_{\mathcal{M}} B(y,z)f(z)dz\right)dy. \end{aligned}$$

In particular, by Hölder inequality with $1/p+1/q=1$,

$$\begin{aligned} &\left|\int_{\mathcal{M}} f(x)d\mu(x)\right| = \left|\int_{\mathcal{M}}\int_{\mathcal{M}}\int_{\mathcal{M}} B(x,y)A(y,z)f(z)d\mu(x)dydz\right| \\ &\le \left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}} A(y,z)f(z)dz\right|^p dy\right\}^{1/p}\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}} B(x,y)d\mu(x)\right|^q dy\right\}^{1/q}. \end{aligned}$$

Finally, when $p=q=2$ the Cauchy inequality reduces to an equality if the functions are proportional. Indeed, if $B\left(x,y\right) = B\left(y,x\right)$ and

$$f(x) = \int_{\mathcal{M}}\int_{\mathcal{M}} B\left(x,y\right)\overline{B\left(y,z\right)}dy\overline{d\mu(z)},$$

then one easily verifies that

$$\int_{\mathcal{M}} f(x)d\mu(x) = \int_{\mathcal{M}}\left|\int_{\mathcal{M}} B\left(x,y\right)d\mu(x)\right|^2 dy,$$
$$\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}} A(x,y)f(y)dy\right|^2 dx\right\}^{1/2} = \left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}} B\left(x,y\right)d\mu(x)\right|^2 dy\right\}^{1/2}.$$

Hence, when $p=q=2$ for this function the inequality in the theorem reduces to an equality. □

In what follows the operators with kernels $A(x,y)$ and $B(x,y)$ will be powers of the Laplace Beltrami operator $(I+\Delta)^{\pm\alpha/2}$.

Definition 2.2. The Sobolev space $W^{\alpha,p}\left(\mathcal{M}\right)$, $-\infty < \alpha < +\infty$ and $1 \le p \le +\infty$, consists of all distributions on $\mathcal{M}$ with $(I+\Delta)^{\alpha/2} f(x)$ in $L^p\left(\mathcal{M}\right)$, that is with

$$\|f\|_{\mathbb{W}^{\alpha,p}} = \left\{\int_{\mathcal{M}}\left|\sum_{\lambda}\left(1+\lambda^2\right)^{\alpha/2}\mathcal{F}f\left(\lambda\right)\varphi_\lambda(x)\right|^p dx\right\}^{1/p} < +\infty.$$

An equivalent definition is the following.

Definition 2.3. Let $B^{\alpha}(x,y)$, $-\infty<\alpha<+\infty$, be the Bessel kernel

$$B^{\alpha}(x,y)=\sum_{\lambda}\left(1+\lambda^{2}\right)^{-\alpha/2}\varphi_{\lambda}(x)\overline{\varphi_{\lambda}(y)}.$$

A distribution $f(x)$ is in the Sobolev space $W^{\alpha,p}\left(\mathcal{M}\right)$ if and only if it is a Bessel potential of a function $g(x)$ in $L^{p}\left(\mathcal{M}\right)$,

$$f(x)=\int_{\mathcal{M}}B^{\alpha}(x,y)g(y)dy.$$

Moreover, $\|f\|_{\mathbb{W}^{\alpha,p}}=\|g\|_{\mathbb{L}^{p}}$.

In particular, when $p=2$,

$$\|f\|_{\mathbb{W}^{\alpha,2}}=\left\{\sum_{\lambda}\left(1+\lambda^{2}\right)^{\alpha}\left|\mathcal{F}f\left(\lambda\right)\right|^{2}\right\}^{1/2}.$$

Another equivalent definition is a localization result: A distribution $f(x)$ is in $\mathbb{W}^{\alpha,p}\left(\mathcal{M}\right)$ if and only if for every smooth function $g(x)$ with support in a local card $x=\psi(y):\mathbb{R}^{d}\rightsquigarrow\mathcal{M}$, the distribution $g(\psi(y))f(\psi(y))$ is in $\mathbb{W}^{\alpha,p}\left(\mathbb{R}^{d}\right)$. In particular, if α is a positive even integer, then $f(x)$ is in $\mathbb{W}^{\alpha,p}\left(\mathcal{M}\right)$ if and only if the p th power of $f(x)$ and of $\Delta^{\alpha/2}f(x)$ are integrable. Moreover, distributions in $\mathbb{W}^{\alpha,p}\left(\mathcal{M}\right)$ with $\alpha>d/p$ are Hölder continuous of degree $\alpha-d/p$. When applied to functions in Sobolev classes, Theorem 2.1 gives the following corollary.

Corollary 2.4. *(1) If $B^{\alpha}(x,y)=\sum_{\lambda}\left(1+\lambda^{2}\right)^{-\alpha/2}\varphi_{\lambda}(x)\overline{\varphi_{\lambda}(y)}$ is the Bessel kernel, if $d\mu(x)$ is a finite measure on $\mathcal{M}$, and if $f(x)$ is a continuous function in $W^{\alpha,p}\left(\mathcal{M}\right)$, with $1\leq p,q\leq+\infty$ and $1/p+1/q=1$, then*

$$\left|\int_{\mathcal{M}}f(x)d\mu(x)\right|\leq\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}B^{\alpha}(x,y)d\mu(x)\right|^{q}dy\right\}^{1/q}\|f\|_{\mathbb{W}^{\alpha,p}}.$$

If $\alpha>d/p$ then the above integrals are well-defined and finite. On the contrary, the spaces $W^{\alpha,p}\left(\mathcal{M}\right)$ with $\alpha\leq d/p$ contain unbounded functions and, if the measure $d\mu(x)$ does not vanish on the set where $f(x)=\infty$, then $\int_{\mathcal{M}}f(x)d\mu(x)$ may diverge.

(2) When $p=q=2$ then the above inequality simplifies,

$$\left|\int_{\mathcal{M}}f(x)d\mu(x)\right|\leq\left\{\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}\left(x,y\right)d\mu(x)\overline{d\mu(y)}\right\}^{1/2}\|f\|_{\mathbb{W}^{\alpha,2}}$$

Equivalently, by the Fourier expansion of the Bessel kernel,

$$\left|\int_{\mathcal{M}}f(x)d\mu(x)\right|\leq\left\{\sum_{\lambda}\left(1+\lambda^{2}\right)^{-\alpha}\left|\mathcal{F}\mu\left(\lambda\right)\right|^{2}\right\}^{1/2}\|f\|_{\mathbb{W}^{\alpha,2}}.$$

Moreover, with $f(x)=\int_{\mathcal{M}}B^{2\alpha}(x,y)\,\overline{d\mu(y)}$ *the above inequalities reduce to equalities.*

(3) If $d\mu(x)=d\nu(x)-dx$ *is the difference between a probability measure* $d\nu(x)$ *and the Riemannian measure* dx*, then*

$$\left|\int_{\mathcal{M}}f(x)d\nu(x)-\int_{\mathcal{M}}f(x)dx\right|\leq\left\{\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,d\nu(x)d\nu(y)-1\right\}^{1/2}\|f\|_{\mathbb{W}^{\alpha,2}}.$$

Equivalently,

$$\left|\int_{\mathcal{M}}f(x)d\nu(x)-\int_{\mathcal{M}}f(x)dx\right|\leq\left\{\sum_{\lambda>0}\left(1+\lambda^{2}\right)^{-\alpha}|\mathcal{F}\nu(\lambda)|^{2}\right\}^{1/2}\|f\|_{\mathbb{W}^{\alpha,2}}.$$

Proof. (1) is an immediate corollary of Theorem 2.1. In order to prove (2), observe that

$$\int_{\mathcal{M}}B^{\alpha}(x,y)B^{\beta}(y,z)dy=B^{\alpha+\beta}(x,z).$$

Moreover, this Bessel kernel is real and symmetric. Hence,

$$\begin{aligned}&\int_{\mathcal{M}}\left|\int_{\mathcal{M}}B^{\alpha}(x,y)d\mu(x)\right|^{2}dy\\&=\int_{\mathcal{M}}\int_{\mathcal{M}}\int_{\mathcal{M}}B^{\alpha}(x,y)B^{\alpha}(z,y)dyd\mu(x)\overline{d\mu(z)}\\&=\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,z)d\mu(x)\overline{d\mu(z)}.\end{aligned}$$

(3) is a corollary of (1) and (2). Indeed, since $B^{2\alpha}(x,y)=B^{2\alpha}(y,x)$ and $\int_{\mathcal{M}}B^{2\alpha}(x,y)\,dy=1$, it follows that

$$\begin{aligned}&\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,(d\nu(x)-dx)\,(d\nu(y)-dy)\\&=\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,d\nu(x)d\nu(y)-\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,d\nu(x)dy\\&\quad-\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,dxd\nu(y)+\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,dxdy\\&=\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,d\nu(x)d\nu(y)-1.\end{aligned}$$

Finally, by Sobolev imbedding theorem, functions in $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with $\alpha>d/p$ are continuous and all the above integrals are well-defined and finite. This also follows from Lemma 2.6 and Remark 3.3 below. □

The above corollary leads to estimate the energy integrals

$$\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}B^{\alpha}(x,y)d\mu(x)\right|^{q}dy\right\}^{1/q},$$

$$\left\{\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,d\nu(x)d\nu(y)-1\right\}^{1/2}=\left\{\sum_{\lambda>0}\left(1+\lambda^{2}\right)^{-\alpha}|\mathcal{F}\nu(\lambda)|^{2}\right\}^{1/2}.$$

By the last formula, the energy attains a minimum if and only if $\mathcal{F}\nu(\lambda) = 0$ for all $\lambda > 0$, and this gives the Riemannian measure dx. The meaning of the corollary is that measures with low energy are close to the Riemannian measure and they give good quadrature rules. In order to give quantitative estimates for the above integrals, one has to collect some properties of the Bessel kernels. The norm of the function $y \rightsquigarrow B^{\alpha}(x,y)$ in $\mathbb{W}^{\gamma,2}(\mathcal{M})$ is

$$\|B^{\alpha}(x,\cdot)\|_{\mathbb{W}^{\gamma,2}} = \left\{\sum_{\lambda} \left(1+\lambda^2\right)^{\gamma-\alpha} |\varphi_\lambda(x)|^2\right\}^{1/2}.$$

By Weyl's estimates on the spectrum of an elliptic operator, see Theorem 17.5.3 in [18], for every $r > 1$ there are approximately cr^d eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$ and $\sum_{\lambda \le r} |\varphi_\lambda(x)|^2 \le cr^d$. It then follows that the norm in $\mathbb{W}^{\gamma,2}(\mathcal{M})$ of $B^{\alpha}(x,y)$ is finite provided that $\gamma < \alpha - d/2$ and, by Sobolev imbedding theorem, it also follows that $B^{\alpha}(x,y)$ is Hölder continuous of degree $\delta < \alpha - d$. Indeed, we shall see that a bit more is true: $B^{\alpha}(x,y)$ is Hölder continuous of degree $\alpha - d$.

Lemma 2.5. *The heat kernel $W(t,x,y) = \sum_{\lambda} \exp(-\lambda^2 t)\,\varphi_\lambda(x)\overline{\varphi_\lambda(y)}$, which is the fundamental solution to the heat equation $\partial/\partial t = -\Delta$ on $\mathbb{R}_+ \times \mathcal{M}$, is symmetric real and positive, $W(t,x,y) = W(t,y,x) > 0$ for every $x, y \in \mathcal{M}$ and $t > 0$. Moreover, for every m and n there exists c such that, if $|x-y|$ denotes the Riemannian distance between x and y, and ∇ the gradient,*

$$\begin{cases} |\nabla^m W(t,x,y)| \le ct^{-(d+m)/2}\left(1+|x-y|/\sqrt{t}\right)^{-n} & \text{if } 0 < t \le 1, \\ |\nabla^m W(t,x,y)| \le c \quad \text{if } 1 \le t < +\infty. \end{cases}$$

Proof. All of this is well known. The idea is that heat has essentially a finite speed of propagation and diffusion in manifolds is comparable to diffusion in Euclidean spaces, at least for small times. The heat kernel in the Euclidean space $\mathbb{R}^d$ is a Gaussian,

$$\begin{aligned} W(t,x,y) &= \int_{\mathbb{R}^d} \exp\left(-4\pi^2 t|\xi|^2\right)\exp\left(2\pi i(x-y)\xi\right)d\xi \\ &= (4\pi t)^{-d/2}\exp\left(-|x-y|^2/4t\right). \end{aligned}$$

By the Poisson summation formula, the heat kernel on the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is the periodized of the kernel in the space,

$$\begin{aligned} &\sum_{k\in\mathbb{Z}^d} \exp\left(-4\pi^2|k|^2 t\right)\exp\left(2\pi i k(x-y)\right) \\ &= \sum_{k\in\mathbb{Z}^d} (4\pi t)^{-d/2}\exp\left(-|x-y-k|^2/4t\right). \end{aligned}$$

When x is close to y and t is small, the main contribution to the sum comes from the term with $k = 0$,

$$W(t,x,y) \approx (4\pi t)^{-d/2}\exp\left(-|x-y|^2/4t\right).$$

The remainder gives a bounded contribution,

$$\left|\sum_{k\in\mathbb{Z}^d-\{0\}}(4\pi t)^{-d/2}\exp\left(-|x-y-k|^2/4t\right)\right|\leq c.$$

Analogous estimates hold for the derivatives. This proves the lemma for the torus. The heat kernel on a compact manifold is similar, in particular it has an asymptotic expansion with euclidean main term. See e.g. [9], Chapter VI. More precisely, by the Minakshisundaram Pleijel recursion formulas, there exist smooth functions $\{u_k(x,y)\}$ such that, if t is small and $|x-y|$ denotes the distance between x and y,

$$W(t,x,y)\approx(4\pi t)^{-d/2}\exp\left(-|x-y|^2/4t\right)\sum_{k=0}^{n}t^k u_k(x,y)+O\left(t^{n+1}\right).$$

On the contrary, $W(t,x,y)\to 1$ when $t\to+\infty$. The estimates on the size of this kernel and its derivatives are a consequence of this asymptotic expansion. The positivity $W(t,x,y)>0$ is a consequence of the maximum principle for heat equation and the symmetry $W(t,x,y)=W(t,y,x)$ follows from this positivity and the eigenfunction expansion. □

Lemma 2.6. *(1) The Bessel kernel $B^\alpha(x,y)$ with $\alpha>0$ is a superposition of heat kernels $W(t,x,y)$:*

$$B^\alpha(x,y)=\Gamma(\alpha/2)^{-1}\int_0^{+\infty}t^{\alpha/2-1}\exp(-t)\,W(t,x,y)\,dt.$$

(2) The Bessel kernel $B^\alpha(x,y)$ with $\alpha>0$ is real and positive for every $x,y\in\mathcal{M}$, and it is smooth in $\{x\neq y\}$. Moreover, for suitable constants $0<a<b$,

$$a|x-y|^{\alpha-d}\leq B^\alpha(x,y)\leq b|x-y|^{\alpha-d}\quad \textit{if } 0<\alpha<d,$$
$$a\log\left(1+|x-y|^{-1}\right)\leq B^\alpha(x,y)\leq b\log\left(1+|x-y|^{-1}\right)\quad \textit{if } \alpha=d,$$
$$a\leq B^\alpha(x,y)\leq b\quad \textit{if } \alpha>d.$$

(3) If $d<\alpha<d+1$, then $B^\alpha(x,y)$ is Hölder continuous of degree $\alpha-d$, that is there exists c such that for every $x,y,z\in\mathcal{M}$,

$$\left|B^\alpha(x,y)-B^\alpha(x,z)\right|\leq c|y-z|^{\alpha-d}.$$

(4) If $d<\alpha<d+2$, then there exists c such that for every $x,y\in\mathcal{M}$,

$$\left|B^\alpha(x,x)-B^\alpha(x,y)\right|\leq c|x-y|^{\alpha-d}.$$

Proof. When the manifold is a torus and the eigenfunctions are exponentials the proof is elementary. The Bessel kernel in the torus $\mathbb{T}^d$ is an even function and it is sum of cosines,

$$\begin{aligned} B^\alpha(x,y) &= \sum_{k\in\mathbb{Z}^d} \left(1+4\pi^2 |k|^2\right)^{-\alpha/2} \exp\left(2\pi ikx\right)\exp\left(-2\pi iky\right) \\ &= \sum_{k\in\mathbb{Z}^d} \left(1+4\pi^2 |k|^2\right)^{-\alpha/2} \cos\left(2\pi k\left(x-y\right)\right). \end{aligned}$$

Hence,

$$\begin{aligned} B^\alpha(x,x) - B^\alpha(x,y) &= 2\sum_{k\in\mathbb{Z}^d} \left(1+4\pi^2 |k|^2\right)^{-\alpha/2} \sin^2\left(\pi k\left(x-y\right)\right) \\ \le 2\pi^2 |x-y|^2 &\sum_{|k|\le|x-y|^{-1}} |k|^2 \left(1+4\pi^2 |k|^2\right)^{-\alpha/2} + 2\sum_{|k|>|x-y|^{-1}} \left(1+4\pi^2 |k|^2\right)^{-\alpha/2} \\ &\le \begin{cases} c\,|x-y|^{\alpha-d} & \text{if } d<\alpha<d+2, \\ c\,|x-y|^2 \log\left(1+|x-y|^{-1}\right) & \text{if } \alpha=d+2, \\ c\,|x-y|^2 & \text{if } \alpha>d+2. \end{cases} \end{aligned}$$

Also observe that the series which defines $B^\alpha(x,x) - B^\alpha(x,y)$ has positive terms and the above inequalities can be reversed. This proves (4) for a torus, and the proof of (3) and (2) is similar. A proof for a generic manifold follows from the representation of Bessel kernels as superposition of heat kernels and the estimates in the previous lemma. In particular, (1) follows from the identity for the Gamma function

$$\left(1+\lambda^2\right)^{-\alpha/2} = \Gamma\left(\alpha/2\right)^{-1}\int_0^{+\infty} t^{\alpha/2-1}\exp\left(-t\left(1+\lambda^2\right)\right)dt.$$

By Lemma 2.5, for every n,

$$0 < W\left(t,x,y\right) \le \begin{cases} ct^{(n-d)/2}\,|x-y|^{-n} & \text{if } 0<t\le|x-y|^2, \\ ct^{-d/2} & \text{if } |x-y|^2\le t\le 1, \\ c & \text{if } t\ge 1. \end{cases}$$

Hence, if $0<\alpha<d$ and $n>d-\alpha$,

$$\begin{aligned} B^\alpha(x,y) &= \Gamma\left(\alpha/2\right)^{-1}\int_0^{+\infty} t^{\alpha/2-1}\exp\left(-t\right)W\left(t,x,y\right)dt \\ \le c\,|x-y|^{-n}\int_0^{|x-y|^2} t^{(\alpha+n-d)/2-1}dt &+ c\int_{|x-y|^2}^1 t^{(\alpha-d)/2-1}dt + \int_1^{+\infty} t^{\alpha/2-1}\exp\left(-t\right)dt \\ &\le c\,|x-y|^{\alpha-d}. \end{aligned}$$

Indeed one can easily see that these inequalities can be reversed. Hence $B^\alpha(x,y) \approx c\,|x-y|^{\alpha-d}$. This proves (2) when $0<\alpha<d$, and the proofs of the cases $\alpha=d$

and $\alpha > d$ are similar. Also the proof of (3) is similar. Write

$$B^\alpha(x,y) - B^\alpha(x,z)$$
$$= \Gamma(\alpha/2)^{-1} \int_0^{+\infty} t^{\alpha/2-1} \exp(-t) \left(W(t,x,y) - W(t,x,z)\right) dt.$$

Then recall that, by Lemma 2.5,

$$|W(t,x,y) - W(t,x,z)| \leq \begin{cases} ct^{-d/2} & \text{if } 0 < t \leq |y-z|^2, \\ ct^{-(d+1)/2}|y-z| & \text{if } |y-z|^2 \leq t \leq 1, \\ c|y-z| & \text{if } t \geq 1. \end{cases}$$

Hence,

$$|B^\alpha(x,y) - B^\alpha(x,z)| \leq c \int_0^{|y-z|^2} t^{(\alpha-d)/2-1} \exp(-t)\, dt$$
$$+c|y-z| \int_{|y-z|^2}^1 t^{(\alpha-d-1)/2-1} \exp(-t)\, dt + c|y-z| \int_1^{+\infty} t^{\alpha/2-1} \exp(-t)\, dt$$
$$\leq c|y-z|^{\alpha-d}.$$

Finally, the estimate for $|B^\alpha(x,x) - B^\alpha(x,y)|$ in (4) is analogous to the previous one, but it holds in a larger range of α. It suffices to observe that $W(t,x,y)$ is stationary at $x = y$ and it satisfies the estimates

$$|W(t,x,x) - W(t,x,y)| \leq \begin{cases} ct^{-d/2} & \text{if } 0 < t \leq |x-y|^2, \\ ct^{-d/2-1}|x-y|^2 & \text{if } |x-y|^2 \leq t \leq 1, \\ c|x-y|^2 & \text{if } t \geq 1. \end{cases}$$

□

The following is Result (1) in the Introduction.

Theorem 2.7. *For every $d/2 < \alpha < d/2+1$ there exists $c > 0$ with the following property: If $\mathcal{M} = U_1 \cup U_2 \cup \ldots \cup U_N$ is a decomposition of $\mathcal{M}$ in disjoint pieces with measure $|U_j| = \omega_j$, then there exists a distribution of points $\{z_j\}_{j=1}^N$ with $z_j \in U_j$ such that*

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \leq c \max_{1\leq j\leq N} \left\{ \text{diameter}(U_j)^\alpha \right\} \|f\|_{\mathbb{W}^{\alpha,2}(\mathcal{M})}.$$

Proof. By Corollary 2.4 (3), with $d\nu(x) = \sum_{j=1}^N \omega_j \delta_{z_j}(x)$,

$$\left| \int_{\mathcal{M}} f(x)d\nu(x) - \int_{\mathcal{M}} f(x)dx \right| \leq \left\{ \sum_{i=1}^N \sum_{j=1}^N \omega_i\omega_j B^{2\alpha}(z_i,z_j) - 1 \right\}^{1/2} \|f\|_{\mathbb{W}^{\alpha,2}}.$$

It suffices to compute the average value of $\sum_{i=1}^{N}\sum_{j=1}^{N}\omega_i\omega_j B^{2\alpha}(z_i,z_j)-1$ on $U_1\times U_2\times \ldots\times U_N$ with respect to the probability measures $\omega_j^{-1}dz_j$ uniformly distributed on U_j. First observe that

$$\left(\prod_{k=1}^{N}\omega_k^{-1}\right)\int_{U_1}\ldots\int_{U_N}dz_1\ldots dz_N=1,$$

$$1=\int_{\mathcal{M}}\int_{\mathcal{M}}B^{2\alpha}(x,y)\,dxdy=\sum_{i=1}^{N}\sum_{j=1}^{N}\int_{U_i}\int_{U_j}B^{2\alpha}(x,y)\,dxdy.$$

Then,

$$\begin{aligned}&\left(\prod_{k=1}^{N}\omega_k^{-1}\right)\int_{U_1}\ldots\int_{U_N}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\omega_i\omega_j B^{2\alpha}(z_i,z_j)-1\right)dz_1\ldots dz_N\\&=\sum_{j}\omega_j\int_{U_j}B^{2\alpha}(z_j,z_j)\,dz_j+\sum\sum_{i\neq j}\int_{U_i}\int_{U_j}B^{2\alpha}(z_i,z_j)\,dz_idz_j\\&-\sum_{j}\int_{U_j}\int_{U_j}B^{2\alpha}(x,y)\,dxdy-\sum\sum_{i\neq j}\int_{U_i}\int_{U_j}B^{2\alpha}(x,y)\,dxdy\\&=\sum_{j=1}^{N}\int_{U_j}\int_{U_j}\left(B^{2\alpha}(x,x)-B^{2\alpha}(x,y)\right)dxdy.\end{aligned}$$

Since, by Lemma 2.6 (4), $|B^{2\alpha}(x,x)-B^{2\alpha}(x,y)|\leq c\,|x-y|^{2\alpha-d}$ when $d<2\alpha<d+2$, and since $\omega_j=|U_j|\leq c\,\text{diameter}\,(U_j)^d$,

$$\begin{aligned}&\sum_{j=1}^{N}\int_{U_j}\int_{U_j}\left|B^{2\alpha}(x,x)-B^{2\alpha}(x,y)\right|dxdy\\&\leq\sum_{j=1}^{N}|U_j|^2\sup\left\{\left|B^{2\alpha}(x,x)-B^{2\alpha}(x,y)\right|,\ x,y\in U_j\right\}\\&\leq c\sum_{j=1}^{N}|U_j|^2\,\text{diameter}\,(U_j)^{2\alpha-d}\leq c\sum_{j=1}^{N}|U_j|\,\text{diameter}\,(U_j)^{2\alpha}.\end{aligned}$$

□

For the next result we shall need estimates for partial sums of Fourier expansions of the Bessel kernels.

Lemma 2.8. *Let $\chi(\lambda)$ be an even smooth function on $-\infty<\lambda<+\infty$ with support in $1/2\leq|\lambda|\leq 2$ and let*

$$P^{\alpha}(r,x,y)=\sum_{\lambda}\chi(\lambda/r)\left(1+\lambda^2\right)^{-\alpha/2}\varphi_\lambda(x)\overline{\varphi_\lambda(y)}.$$

Then for every $n > 0$ there exists c such that for every $r > 1$ and $x, y \in M$,

$$|P^\alpha(r,x,y)| \leq cr^{d-\alpha}\left(1+r\left|x-y\right|\right)^{-n}.$$

Proof. The numerology behind this estimate is quite simple. The approximation of the Bessel kernel $B^\alpha(x,y)$ by linear combinations of eigenfunctions with eigenvalues $\lambda^2 < r^2$ is localized and only points x and y with $|x-y| \leq 1/r$ really matter. In particular, since $B^\alpha(x,y)$ is smooth away from the diagonal, at distance $|x-y| \leq 1/r$ the approximation is rough, but at distance $|x-y| \geq 1/r$ it is quite good. The analogue of $P^\alpha(r,x,y)$ in the Euclidean setting is the kernel

$$\begin{aligned} Q\left(r,x-y\right) &= \int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|/r\right)\left(1+4\pi^2\left|\xi\right|^2\right)^{-\alpha/2}\exp\left(2\pi i\left(x-y\right)\xi\right)d\xi \\ &= r^d\int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\exp\left(2\pi i r\left(x-y\right)\xi\right)d\xi. \end{aligned}$$

Since $\chi\left(2\pi\left|\xi\right|\right)$ has support in $1/2 \leq 2\pi\left|\xi\right| \leq 2$, for every r and $x, y \in \mathbb{R}^d$ one has

$$\begin{aligned} &\left|r^d\int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\exp\left(2\pi i r\left(x-y\right)\xi\right)d\xi\right| \\ &\qquad \leq r^{d-\alpha}\int_{\mathbb{R}^d}\left(2\pi\left|\xi\right|\right)^{-\alpha}\left|\chi\left(2\pi\left|\xi\right|\right)\right|d\xi \leq cr^{d-\alpha}. \end{aligned}$$

This estimate can be improved in the range $|x-y| \geq 1/r$. Indeed, an integration by parts gives

$$\begin{aligned} &r^d\int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\exp\left(2\pi i r\left(x-y\right)\xi\right)d\xi \\ &= r^d\int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\Delta_\xi^n\left(\left(4\pi^2 r^2\left|x-y\right|^2\right)^{-n}\exp\left(2\pi i r\left(x-y\right)\xi\right)\right)d\xi \\ &= r^d\left(4\pi^2 r^2\left|x-y\right|^2\right)^{-n}\int_{\mathbb{R}^d}\exp\left(2\pi i r\left(x-y\right)\xi\right)\Delta_\xi^n\left(\chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\right)d\xi. \end{aligned}$$

Hence,

$$\begin{aligned} &\left|r^d\int_{\mathbb{R}^d} \chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\exp\left(2\pi i r\left(x-y\right)\xi\right)d\xi\right| \\ &\leq r^d\left(4\pi^2 r^2\left|x-y\right|^2\right)^{-n}\int_{\mathbb{R}^d}\left|\Delta_\xi^n\left(\chi\left(2\pi\left|\xi\right|\right)\left(1+4\pi^2 r^2\left|\xi\right|^2\right)^{-\alpha/2}\right)\right|d\xi \\ &\leq cr^{d-\alpha-2n}\left|x-y\right|^{-2n}. \end{aligned}$$

Now it suffices to transfer these estimates from the Euclidean space to the manifold. For the torus, this can be done via the Poisson summation formula. If $Q\left(r,x-y\right)$ is the truncated Bessel kernel in $\mathbb{R}^d$ defined above, then the truncated Bessel kernel in $\mathbb{T}^d$ is

$$\sum_{k\in\mathbb{Z}^d}\chi\left(2\pi\left|k\right|/r\right)\left(1+4\pi^2\left|h\right|^2\right)^{-\alpha/2}\exp\left(2\pi i k\left(x-y\right)\right) = \sum_{k\in\mathbb{Z}^d} Q\left(r,x-y+k\right).$$

When $|x_j - y_j| \leq 1/2$, the main term in the last sum is the one with $k = 0$, while the contribution of terms with $k \neq 0$ is negligible,

$$|Q(r, x-y)| \leq cr^{d-\alpha}(1 + r|x-y|)^{-n},$$
$$\sum_{k \in \mathbb{Z}^d - \{0\}} |Q(r, x-y-k)| \leq cr^{d-\alpha-n}.$$

Finally, the estimate for the truncated Bessel kernel on a generic manifold can be obtained by transference from $\mathbb{R}^d$ via pseudodifferential techniques. For more details, see e.g. [27] Chapter XII, or [5]. □

The following is a result on the homogeneity of measures which appear in quadrature rules and it gives sharp estimates of the discrepancy of such measures. Similar estimates on spheres are in [2].

Lemma 2.9. *Assume that $d\nu(x)$ is a probability measure on $\mathcal{M}$ with the property that for every eigenfunction $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$,*

$$\int_{\mathcal{M}} \varphi_\lambda(x) d\nu(x) = \int_{\mathcal{M}} \varphi_\lambda(x) dx.$$

Then for every n there exists c, which may depend on n and $\mathcal{M}$, but is independent of r and $d\nu(x)$, such that for every measurable set Ω in $\mathcal{M}$,

$$\left| \int_\Omega d\nu(x) - \int_\Omega dx \right| \leq c \int_{\mathcal{M}} (1 + r\text{distance}\{x, \partial\Omega\})^{-n} dx.$$

In particular, the discrepancy between the measures $d\nu(x)$ and dx with respect to balls $\{|x-y| \leq s\}$ is dominated by

$$\left| \int_{\{|x-y|\leq s\}} d\nu(x) - \int_{\{|x-y|\leq s\}} dx \right| \leq \begin{cases} cr^{-d} & \textit{if } s \leq 1/r, \\ cr^{-1}s^{d-1} & \textit{if } s \geq 1/r. \end{cases}$$

Proof. It is proved in [11] that given n, there exists c such that for every measurable set Ω in $\mathcal{M}$ and every $r > 0$ there exist two linear combinations of eigenfunctions $A(x) = \sum_{\lambda<r} a(\lambda)\varphi_\lambda(x)$ and $B(x) = \sum_{\lambda<r} b(\lambda)\varphi_\lambda(x)$ which approximate the characteristic function $\chi_\Omega(x)$ from above and below,

$$A(x) \leq \chi_\Omega(x) \leq B(x), \quad B(x) - A(x) \leq c(1 + r\text{distance}\{x, \partial\Omega\})^{-n}.$$

In particular, the properties of the function $A(x)$ and of the measure $d\nu(x)$ give

$$\int_\Omega d\nu(x) \geq \int_{\mathcal{M}} A(x) d\nu(x) = \int_{\mathcal{M}} A(x) dx$$
$$\geq \int_{\mathcal{M}} \chi_\Omega(x) dx - c \int_{\mathcal{M}} (1 + r\text{distance}\{x, \partial\Omega\})^{-n} dx.$$

Similarly, by the properties of $B(x)$ and $d\nu(x)$,

$$\int_{\Omega} d\nu(x) \leq \int_{\mathcal{M}} B(x)d\nu(x) = \int_{\mathcal{M}} B(x)dx$$
$$\leq \int_{\mathcal{M}} \chi_{\Omega}(x)dx + c\int_{\mathcal{M}} (1 + r\text{distance}\{x, \partial\Omega\})^{-n} dx.$$

□

Lemma 2.10. *Assume that $d\nu(x)$ is a probability measure on $\mathcal{M}$ which gives an exact quadrature for all eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$,*

$$\int_{\mathcal{M}} \varphi_\lambda(x)d\nu(x) = \int_{\mathcal{M}} \varphi_\lambda(x)dx.$$

If $1 \leq q \leq +\infty$ and $\alpha > d(1 - 1/q)$, then there exists c, which may depend on q, α, $\mathcal{M}$, but is independent of r and $d\nu(x)$, such that

$$\left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} B^\alpha(x, y)d\nu(x) - 1\right|^q dy\right\}^{1/q} \leq cr^{-\alpha}.$$

Proof. Let $\chi(\lambda)$ be an even smooth function on $-\infty < \lambda < +\infty$ with support in $1/2 \leq |\lambda| \leq 2$ with the property that $\sum_{j=-\infty}^{+\infty} \chi(2^{-j}\lambda) = 1$ for every $\lambda \neq 0$. Also, let

$$P^\alpha(s, x, y) = \sum_\lambda \chi(\lambda/s)\left(1 + \lambda^2\right)^{-\alpha/2} \varphi_\lambda(x)\overline{\varphi_\lambda(y)}.$$

Hence, $B^\alpha(x, y) = 1 + \sum_{j=-\infty}^{+\infty} P^\alpha(2^j, x, y)$. Since $d\nu(x)$ annihilates all eigenfunctions with $0 < \lambda < r$, it also annihilates all $P^\alpha(2^j, x, y)$ with $2^j \leq r/2$ and this gives

$$\int_{\mathcal{M}} B^\alpha(x, y)d\nu(x) - 1 = \int_{\mathcal{M}} \left(\sum_{2^j > r/2} P^\alpha(2^j, x, y)\right) d\nu(x).$$

When $q = 1$, by Lemma 2.8,

$$\int_{\mathcal{M}} \left|\int_{\mathcal{M}} P^\alpha(s, x, y)d\nu(x)\right| dy$$
$$\leq cs^{d-\alpha} \int_{\mathcal{M}}\int_{\mathcal{M}} (1 + s|x - y|)^{-n} d\nu(x)dy$$
$$\leq cs^{-\alpha} \sup_{x\in\mathcal{M}} \left\{\int_{\mathcal{M}} s^d (1 + s|x - y|)^{-n} dy\right\} \leq cs^{-\alpha}.$$

When $q=+\infty$ and $s\geq r$ and $n>d$, by Lemma 2.8 and Lemma 2.9,

$$\begin{aligned}
&\sup_{y\in\mathcal{M}}\left\{\left|\int_{\mathcal{M}}P^{\alpha}(s,x,y)d\nu(x)\right|\right\}\\
&\leq cs^{d-\alpha}\sup_{y\in\mathcal{M}}\left\{\int_{\mathcal{M}}(1+s\left|x-y\right|)^{-n}d\nu(x)\right\}\\
&\leq cs^{d-\alpha}\sup_{y\in\mathcal{M}}\left\{\int_{\{|x-y|\leq 1/r\}}d\nu(x)\right\}\\
&+cs^{d-\alpha}\sup_{y\in\mathcal{M}}\left\{\sum_{j=0}^{+\infty}\left(2^{j}s/r\right)^{-n}\int_{\{|x-y|\leq 2^{j}/r\}}d\nu(x)\right\}\\
&\leq cs^{d-\alpha}r^{-d}+cs^{d-\alpha-n}r^{n-d}\leq cs^{d-\alpha}r^{-d}.
\end{aligned}$$

Hence, when $s\geq r$ and $1<q<+\infty$, by interpolation between 1 and $+\infty$,

$$\begin{aligned}
&\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}P^{\alpha}(s,x,y)d\nu(x)\right|^{q}dy\right\}^{1/q}\\
&\leq\sup_{y\in\mathcal{M}}\left\{\left|\int_{\mathcal{M}}P^{\alpha}(s,x,y)d\nu(x)\right|\right\}^{(q-1)/q}\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}P^{\alpha}(s,x,y)d\nu(x)\right|dy\right\}^{1/q}\\
&\leq cs^{d(1-1/q)-\alpha}r^{-d(1-1/q)}.
\end{aligned}$$

When $\alpha>d(1-1/q)$ these estimates sum to

$$\begin{aligned}
&\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}B^{\alpha}(x,y)d\nu(x)-1\right|^{q}dy\right\}\\
&\leq\sum_{2^{j}>r/2}\left\{\int_{\mathcal{M}}\left|\int_{\mathcal{M}}P^{\alpha}(2^{j},x,y)d\nu(x)\right|^{q}dy\right\}^{1/q}\\
&\leq cr^{-d(1-1/q)}\sum_{2^{j}>r/2}2^{j(d(1-1/q)-\alpha)}\leq cr^{-\alpha}.
\end{aligned}$$

□

Finally, the existence of exact quadrature rules associated to any system of continuous functions is a simple result in functional analysis, or in convex geometry. See Theorem 3.1.1 in [26], or [25], or [8] for explicit constructions on spheres.

Lemma 2.11. *Given any number $\varphi_1(x)$, $\varphi_2(x)$,..., $\varphi_n(x)$ of continuous functions on $\mathcal{M}$, there exist points $\{z_j\}_{j=1}^{N}$ in $\mathcal{M}$ and positive weights $\{\omega_j\}_{j=1}^{N}$ with $\sum_{j=1}^{N}\omega_j=1$, such that for every $\varphi_i(x)$,*

$$\int_{\mathcal{M}}\varphi_i(x)dx=\sum_{j=1}^{N}\omega_j\varphi_i\left(z_j\right).$$

If the functions $\varphi_j(x)$ are real one can choose $N\leq n+1$, and if these functions are complex one can choose $N\leq 2n+1$.

Proof. Define

$$\Phi(x) = (\varphi_1(x), \varphi_2(x), ..., \varphi_n(x)),$$
$$E = \int_{\mathcal{M}} \Phi(x)dx = \left(\int_{\mathcal{M}} \varphi_1(x)dx, \int_{\mathcal{M}} \varphi_2(x)dx, ..., \int_{\mathcal{M}} \varphi_n(x)dx \right).$$

If all functions $\varphi_i(x)$ are real valued, then $\Phi(x)$ and E are vectors in $\mathbb{R}^n$. If the $\varphi_i(x)$ are complex, then $\Phi(x)$ and E can be seen as vectors in $\mathbb{R}^{2n}$. Moreover, E is in the convex hull of the vectors $\Phi(x)$ with $x \in \mathcal{M}$. It then follows from Caratheodory's theorem that E is also a convex combination of at most $n+1$ vectors $\Phi(x)$ in the real case, or $2n+1$ in the complex case, $E = \sum_{j=1}^{N} \omega_j \Phi(z_j)$ with $\omega_j \geq 0$ and $\sum_{j=1}^{N} \omega_j = 1$. □

The following is Result (2) in the Introduction.

Theorem 2.12. *Assume that the probability measure $d\nu(x)$ on $\mathcal{M}$ gives an exact quadrature for all eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$,*

$$\int_{\mathcal{M}} \varphi_\lambda(x)d\nu(x) = \int_{\mathcal{M}} \varphi_\lambda(x)dx.$$

Then, for some constant c independent of $d\nu(x)$ and r and for every function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$ with $1 \leq p \leq +\infty$ and $\alpha > d/p$,

$$\left| \int_{\mathcal{M}} f(x)d\nu(x) - \int_{\mathcal{M}} f(x)dx \right| \leq cr^{-\alpha} \|f\|_{\mathbb{W}^{\alpha,p}}.$$

Proof. By Corollary 2.4 (1) with $d\mu(x) = d\nu(x) - dx$,

$$\left| \int_{\mathcal{M}} f(x)d\mu(x) \right| \leq \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x,y)d\nu(x) - 1 \right|^q dy \right\}^{1/q} \|f\|_{\mathbb{W}^{\alpha,p}}.$$

By the assumption $\int_{\mathcal{M}} \varphi_\lambda(x)d\mu(x) = 0$ for every $\lambda < r$, and Lemma 2.10,

$$\left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x,y)d\nu(x) - 1 \right|^q dy \right\}^{1/q} \leq cr^{-\alpha}.$$

□

The above theorem has as corollary Result (3) in the Introduction.

Corollary 2.13. *If $1 \leq p \leq +\infty$ and $\alpha > d/p$, then there exists $c > 0$ with the property that for every N there exist sequences of points $\{z_j\}_{j=1}^N$ and non negative weights $\{\omega_j\}_{j=1}^N$, such that for every function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$,*

$$\left| \sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \leq cN^{-\alpha/d} \|f\|_{\mathbb{W}^{\alpha,p}}.$$

Proof. By Weyl's estimates on the spectrum of an elliptic operator, see Theorem 17.5.3 in [18], for a given r there are approximately cr^d eigenfunctions $\varphi_\lambda(x)$ with $\lambda < r$. The corollary then follows from Lemma 2.11 and Theorem 2.12 with $r = N^{1/d}$. □

This corollary for the sphere is contained in [17]. Finally, easy examples show that the above estimates for the error in approximate quadrature are, up to constants, best possible. Again, see [15] for the case of the sphere. In particular, the following is Result (4) in the Introduction.

Theorem 2.14. *For every $1 \le p \le +\infty$ and $\alpha > 0$ there exists $c > 0$ with the following property: For every distribution of points $\{z_j\}_{j=1}^N$ there exists a function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$ which vanishes in a neighborhood of these points and satisfies*

$$\|f\|_{\mathbb{W}^{\alpha,p}} \le cN^{\alpha/d}, \quad \int_{\mathcal{M}} f(x)dx = 1.$$

As a consequence, for every distribution of points $\{z_j\}_{j=1}^N$ and complex weights $\{\omega_j\}_{j=1}^N$, there exists a function $f(x)$ with

$$\left|\sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx\right| \ge cN^{-\alpha/d} \|f\|_{\mathbb{W}^{\alpha,p}}.$$

Proof. If ε is small, then one can choose $2N$ disjoint balls in $\mathcal{M}$ with diameters $\varepsilon N^{-1/d}$ and, given N points $\{z_j\}$, at least N balls have no points inside. On each empty ball construct a bump function $\psi_j(x)$ supported on it with

$$\|\psi_j\|_{\mathbb{W}^{\alpha,p}} \le cN^{\alpha/d-1/p}, \quad \int_{\mathcal{M}} \psi_j(x)dx = N^{-1}.$$

Define $f(x) = \sum \psi_j(x)$. Then,

$$\|f\|_{\mathbb{W}^{\alpha,p}} \le cN^{\alpha/d}, \quad \int_{\mathcal{M}} f(x)dx = 1.$$

The estimate of the $\mathbb{L}^p(\mathcal{M})$ norms of $(I+\Delta)^{\alpha/2}\psi_j(x)$ and $(I+\Delta)^{\alpha/2} f(x)$ when $\alpha/2$ is an integer follows from the fact that $(I+\Delta)^{\alpha/2}$ is a differential operator and the terms $(I+\Delta)^{\alpha/2}\psi_j(x)$ have disjoint supports. When $\alpha/2$ is not an integer the estimate follows by complex interpolation. Anyhow, the case $p = 2$ is elementary. If $\delta > 1$ and $\alpha\delta$ is an integer,

$$\|f\|_{\mathbb{W}^{\alpha,2}} = \left\{\sum_\lambda \left(1+\lambda^2\right)^\alpha |\mathcal{F}f(\lambda)|^2\right\}^{1/2}$$
$$\le \left\{\sum_\lambda |\mathcal{F}f(\lambda)|^2\right\}^{(1-1/\delta)/2} \left\{\sum_\lambda \left(1+\lambda^2\right)^{\alpha\delta} |\mathcal{F}f(\lambda)|^2\right\}^{1/2\delta} \le cN^{\alpha/d}.$$

□

3 Further results

The following is Result (5) in the Introduction and it states that a quadrature rule which gives an optimal error in the Sobolev space $\mathbb{W}^{\alpha,2}(\mathcal{M})$ is also optimal in all spaces $\mathbb{W}^{\beta,2}(\mathcal{M})$ with $d/2 < \beta < \alpha$.

Theorem 3.1. *If $d\nu(x)$ is a probability measure on $\mathcal{M}$, then the norm of the measure $d\nu(x) - dx$ as a linear functional on $W^{\alpha,2}(\mathcal{M})$ decreases as α increases. Moreover, if the norm of $d\nu(x) - dx$ on $W^{\alpha,2}(\mathcal{M})$ is $r^{-\alpha}$,*

$$\left|\int_{\mathcal{M}} f(x)d\nu(x) - \int_{\mathcal{M}} f(x)dx\right| \leq r^{-\alpha} \|f\|_{\mathbb{W}^{\alpha,2}},$$

then for every $d/2 < \beta < \alpha$ there exists a constant c which may depend on α, β, $\mathcal{M}$, but is independent of r and $d\nu(x)$, such that

$$\left|\int_{\mathcal{M}} f(x)d\nu(x) - \int_{\mathcal{M}} f(x)dx\right| \leq cr^{-\beta} \|f\|_{\mathbb{W}^{\beta,2}}.$$

Proof. By Corollary 2.4 (2) the norm of the measure $d\nu(x) - dx$ as a linear functional on $\mathbb{W}^{\alpha,2}(\mathcal{M})$ is

$$\left\{\int_{\mathcal{M}}\int_{\mathcal{M}} B^{2\alpha}(x,y)\, d\nu(x)d\nu(y) - 1\right\}^{1/2} = \left\{\sum_{\lambda>0} \left(1+\lambda^2\right)^{-\alpha} |\mathcal{F}\nu(\lambda)|^2\right\}^{1/2}.$$

Since $(1+\lambda^2)^{-\alpha} \leq (1+\lambda^2)^{-\beta}$ when $\beta < \alpha$, it follows that this norm is a decreasing function of α. Write $d\nu(x) - dx = d\mu(x)$. By Lemma 2.6 (1), the norm of the functional $\int_{\mathcal{M}} f(x)d\mu(x)$ on $\mathbb{W}^{\alpha,2}(\mathcal{M})$ can be written as

$$\begin{aligned}&\left\{\int_{\mathcal{M}}\int_{\mathcal{M}} B^{2\alpha}(x,y)\, d\mu(x)\overline{d\mu(y)}\right\}^{1/2} \\ &= \left\{\Gamma(\alpha)^{-1}\int_0^{+\infty} t^{\alpha-1}\exp(-t)\left(\int_{\mathcal{M}}\int_{\mathcal{M}} W(t,x,y)\, d\mu(x)\overline{d\mu(y)}\right)dt\right\}^{1/2}.\end{aligned}$$

Assuming that this norm is $r^{-\alpha}$, one has to show that the corresponding expression with β instead of α is at most $cr^{-\beta}$. Since $\beta < \alpha$, the integral over $1 \leq t < +\infty$ satisfies the estimate

$$\begin{aligned}&\int_1^{+\infty} t^{\beta-1}\exp(-t)\left(\int_{\mathcal{M}}\int_{\mathcal{M}} W(t,x,y)\, d\mu(x)\overline{d\mu(y)}\right)dt \\ &\leq \int_1^{+\infty} t^{\alpha-1}\exp(-t)\left(\int_{\mathcal{M}}\int_{\mathcal{M}} W(t,x,y)\, d\mu(x)\overline{d\mu(y)}\right)dt \\ &\leq \Gamma(\alpha)\, r^{-2\alpha}.\end{aligned}$$

Similarly, since $\beta < \alpha$ the integral over $r^{-2} \le t \le 1$ satisfies the estimate

$$\int_{r^{-2}}^{1} t^{\beta-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t,x,y) \, d\mu(x) \overline{d\mu(y)} \right) dt$$
$$\le r^{2\alpha-2\beta} \int_{r^{-2}}^{1} t^{\alpha-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t,x,y) \, d\mu(x) \overline{d\mu(y)} \right) dt$$
$$\le \Gamma(\alpha) \, r^{-2\beta}.$$

Finally, by the Gaussian estimate on the heat kernel in the proof of Lemma 2.5, if $0 < t < r^{-2}$ then

$$t^{d/2} W(t,x,y) \le c r^{-d} W\left(r^{-2},x,y\right).$$

It then follows that if $\beta > d/2$ the integral over $0 \le t \le r^{-2}$ satisfies the estimate

$$\int_{0}^{r^{-2}} t^{\beta-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t,x,y) \, d\mu(x) \overline{d\mu(y)} \right) dt$$
$$\le c r^{-2\beta} \int_{\mathcal{M}} \int_{\mathcal{M}} W\left(r^{-2},x,y\right) d|\mu|(x) d|\mu|(y).$$

It remains to show that the last double integral is uniformly bounded in r. Since $d|\mu|(x) = d\nu(x) + dx$ and since $\int_{\mathcal{M}} W(r^{-2},x,y)\, dx = 1$, replacing $d|\mu|(x)$ with $d\mu(x)$ it suffices to show that

$$\int_{\mathcal{M}} \int_{\mathcal{M}} W\left(r^{-2},x,y\right) d\mu(x) d\mu(y) \le c.$$

By the assumption on $d\mu(x)$ and the eigenfunction expansion of $W(r^{-2},x,y)$,

$$\int_{\mathcal{M}} \int_{\mathcal{M}} W\left(r^{-2},x,y\right) d\mu(x) d\mu(y)$$
$$\le r^{-\alpha} \left\| \int_{\mathcal{M}} W\left(r^{-2},x,y\right) d\mu(y) \right\|_{\mathbb{W}^{\alpha,2}}$$
$$= r^{-\alpha} \left\{ \sum_{\lambda} \left(1+\lambda^2\right)^{\alpha} \exp\left(-\left(\lambda/r\right)^2\right) |\mathcal{F}\mu(\lambda)|^2 \right\}^{1/2}$$
$$\le r^{-\alpha} \left\{ \sum_{\lambda} \left(1+\lambda^2\right)^{-\alpha} |\mathcal{F}\mu(\lambda)|^2 \right\}^{1/2} \sup_{\lambda} \left\{ \left(1+\lambda^2\right)^{\alpha} \exp\left(-\left(\lambda/r\right)^2/2\right) \right\}.$$

Finally, the last sum with $\{\mathcal{F}\mu(\lambda)\}$ is the norm of the measure $d\mu(x)$ as functional on $\mathbb{W}^{\alpha,2}(\mathcal{M})$, hence by assumption it is $r^{-\alpha}$, and the last supremum is dominated by $r^{2\alpha}$. □

As we said, the above result is only one way, from α to $\beta < \alpha$. If the norm of $d\nu(x) - dx$ on $\mathbb{W}^{\alpha,p}(\mathcal{M})$ is $r^{-\alpha}$ and if $\beta > \alpha$, then one cannot conclude that the norm of $d\nu(x) - dx$ on $\mathbb{W}^{\beta,p}(\mathcal{M})$ is at most $cr^{-\beta}$. As a counterexample, it suffices to perturb a good quadrature rule with nodes $\{z_j\}_{j=1}^{N}$ and weights $\{\omega_j\}_{j=1}^{N}$ by moving the last point z_N into a new point t_N, so that the new quadrature differs from the old one by the quantity $\omega_N |f(z_N) - f(t_N)|$. If $\alpha > d/p + 1$ then the function f is differentiable and $\omega_N |f(z_N) - f(t_N)| \approx \omega_N |z_N - t_N|$. Then, by choosing $|z_N - t_N| = r^{-\alpha}/\omega_N$ one obtains a quadrature rule which gives an error $\approx r^{-\alpha}$ in all spaces $\mathbb{W}^{\beta,p}(\mathcal{M})$ with $\beta > \alpha$. The counterexample when $d/p < \alpha \leq d/p + 1$ is slightly more complicated but similar.

In all the above results, the accuracy in a quadrature rule has been estimated in terms of the energy of a measure. It is also possible to estimate this accuracy in terms of a geometric discrepancy. The Bessel kernel can be decomposed as superposition of characteristic functions,

$$B^\alpha(x,y) = \int_0^{+\infty} \chi_{\{B^\alpha(x,y)>t\}}(x)dt.$$

If $1 \leq p, q \leq +\infty$ and $1/p + 1/q = 1$, by Corollary 2.4 and Minkowski inequality, the following Koksma Hlawka type inequality holds:

$$\left|\int_{\mathcal{M}} f(x)d\mu(x)\right| \leq \|f\|_{\mathbb{W}^{\alpha,p}} \left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} B^\alpha(x,y)d\mu(x)\right|^q dy\right\}^{1/q}$$
$$\leq \|f\|_{\mathbb{W}^{\alpha,p}} \int_0^{+\infty} \left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} \chi_{\{B^\alpha(x,y)>t\}}(x)d\mu(x)\right|^q dy\right\}^{1/q} dt.$$

The quantity $\left|\int_{\mathcal{M}} \chi_{\{B^\alpha(x,y)>t\}}(x)d\mu(x)\right|$ is the discrepancy of the measure $d\mu(x)$ with respect to the level sets $\{B^\alpha(x,y) > t\}$. It can be proved that, for specific measures and at least in the range $0 < \alpha < 1$, the above estimates are sharp and they can lead to optimal quadrature rules. In particular, the following is Result (6) in the Introduction.

Theorem 3.2. *Denote by $\delta(t)$ the diameter of the level sets of the Bessel kernel $\{B^\alpha(x,y) > t\}$ and assume that there exists $r \geq 1$ such that the discrepancy of the measure $d\mu(x)$ with respect to $\{B^\alpha(x,y) > t\}$ satisfies the estimates*

$$\left|\int_{\mathcal{M}} \chi_{\{B^\alpha(x,y)>t\}}(x)d\mu(x)\right| \leq \begin{cases} r^{-d} & \text{if } \delta(t) \leq 1/r, \\ r^{-1}\delta(t)^{d-1} & \text{if } \delta(t) \geq 1/r. \end{cases}$$

Also assume that $1 \leq p \leq +\infty$ and $\alpha > d/p$. Then there exists a constant c, which may depend on α and p and on the total variation of the measure $|\mu|(\mathcal{M})$, but is independent of r, such that

$$\left|\int_{\mathcal{M}} f(x)d\mu(x)\right| \leq \begin{cases} cr^{-\alpha}\|f\|_{\mathbb{W}^{\alpha,p}} & \text{if } 0 < \alpha < 1, \\ cr^{-1}\log(1+r)\|f\|_{\mathbb{W}^{\alpha,p}} & \text{if } \alpha = 1, \\ cr^{-1}\|f\|_{\mathbb{W}^{\alpha,p}} & \text{if } \alpha > 1. \end{cases}$$

Proof. Observe that the above hypotheses on the discrepancy match the estimates in Lemma 2.9. Indeed, by this lemma, the measures $d\nu(x)$ which give exact quadrature for eigenfunctions with eigenvalues $\lambda^2 < r^2$ have discrepancy

$$\left|\int_{\{|x-y|\leq s\}} d\nu(x) - \int_{\{|x-y|\leq s\}} dx\right| \leq \begin{cases} cr^{-d} & \text{if } s \leq 1/r, \\ cr^{-1}s^{d-1} & \text{if } s \geq 1/r. \end{cases}$$

Actually, these estimates hold not only for balls $\{|x-y| \leq s\}$, but also for sets with boundaries with finite $d-1$ dimensional Minkowski measure, such as the level sets $\{B^{\alpha}(x,y) > t\}$. Also observe that these estimates are natural, since the discrepancy of large sets is qualitatively different from the one of small sets. If $1 \leq p,q \leq +\infty$ and $1/p + 1/q = 1$, by Corollary 2.4 and Minkowski inequality,

$$\left|\int_{\mathcal{M}} f(x)d\mu(x)\right| \leq \|f\|_{\mathbb{W}^{\alpha,p}} \int_0^{+\infty} \left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|^q dy\right\}^{1/q} dt.$$

By Lemma 2.6, when $0 < \alpha < d$ then $B^{\alpha}(x,y) \approx |x-y|^{\alpha-d}$, the level sets $\{B^{\alpha}(x,y) > t\}$ have diameters $\delta(t) \approx \min\left\{1, t^{1/(\alpha-d)}\right\}$ and the boundaries $\{B^{\alpha}(x,y) = t\}$ have surface measure of the order of $\delta(t)^{d-1} \approx \min\left\{1, t^{(d-1)/(\alpha-d)}\right\}$. Hence the estimate of the discrepancy of small level sets with $t \geq r^{d-\alpha}$ gives

$$\begin{aligned}
&\left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|^q dy\right\}^{1/q} \\
&\leq \sup_{y\in\mathcal{M}} \left\{\left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|\right\}^{(q-1)/q} \left\{\int_{\mathcal{M}}\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d|\mu|(x)dy\right\}^{1/q} \\
&\leq \sup_{y\in\mathcal{M}} \left\{\left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|\right\}^{(q-1)/q} \left\{c|\mu|(\mathcal{M})t^{d/(\alpha-d)}\right\}^{1/q} \\
&\leq cr^{-d(q-1)/q}t^{d/q(\alpha-d)}.
\end{aligned}$$

Hence, if $\alpha > d/p$ the integral over $r^{d-\alpha} \leq t < +\infty$ satisfies the inequality

$$\begin{aligned}
&\int_{r^{d-\alpha}}^{+\infty} \left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|^q dy\right\}^{1/q} dt \\
&\leq cr^{-d(q-1)/q} \int_{r^{d-\alpha}}^{+\infty} t^{d/q(\alpha-d)}dt \leq cr^{-\alpha}.
\end{aligned}$$

Similarly, the integral over $0 \leq t \leq r^{d-\alpha}$, that is the discrepancy of large level sets, satisfies the inequality

$$\begin{aligned}
&\int_0^{r^{d-\alpha}} \left\{\int_{\mathcal{M}} \left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right|^q dy\right\}^{1/q} dt \\
&\leq r^{-1}\int_0^{r^{d-\alpha}} \min\left\{1, t^{(d-1)/(\alpha-d)}\right\} dt \leq \begin{cases} cr^{-\alpha} & \text{if } 0 < \alpha < 1, \\ cr^{-1}\log(1+r) & \text{if } \alpha = 1, \\ cr^{-1} & \text{if } \alpha > 1. \end{cases}
\end{aligned}$$

The proof in the case $\alpha = d$ is similar and it follows from the estimate $B^{\alpha}(x,y) \approx -\log(|x-y|)$. The proof in the case $\alpha > d$ is even simpler, since in this case $B^{\alpha}(x,y)$ is bounded and it suffices to integrate on $0 \leq t \leq \sup_{x,y\in\mathcal{M}} B^{\alpha}(x,y)$ the inequality $\left|\int_{\mathcal{M}} \chi_{\{B^{\alpha}(x,y)>t\}}(x)d\mu(x)\right| \leq cr^{-1}$. □

In particular, it follows from Lemma 2.9, Theorem 2.12, Theorem 2.14, that, at least in the range $0 < \alpha < 1$, Theorem 3.1 gives an optimal quadrature. We conclude with a series of remarks.

Remark 3.3. As we said, the assumption $\alpha > d/2$ with $p = 2$ in Theorem 2.7, or $\alpha > d/p$ with $1 \leq p \leq +\infty$ in Theorem 2.12, guarantees the boundedness and continuity of $f(x)$, otherwise $f(z_j)$ may be not defined. This follows from the Sobolev imbedding theorem. Indeed, the imbedding is an easy corollary of Lemma 2.6. A function $f(x)$ is in the Sobolev space $\mathbb{W}^{\alpha,p}(\mathcal{M})$ if and only if there exists a function $g(x)$ in $\mathbb{L}^p(\mathcal{M})$ with

$$f(x) = \int_{\mathcal{M}} B^{\alpha}(x,y)g(y)dy.$$

When $1 \leq p,q \leq +\infty$, $1/p + 1/q = 1$, $d/p < \alpha < d$, then $B^{\alpha}(x,y) \leq c\,|x-y|^{\alpha-d}$ is in $\mathbb{L}^q(\mathcal{M})$ and this implies that distributions in the Sobolev space $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are continuous functions. Indeed they are also Hölder continuous of order $\alpha - d/p$.

Remark 3.4. When the manifold is a Lie group or a homogeneous space, one can restate Theorem 2.1 in terms of convolutions. In the particular case of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, let

$$A(x) = \sum_{k\in\mathbb{Z}^d} \psi(k)\exp(2\pi ikx), \quad B(x) = \sum_{k\in\mathbb{Z}^d} \psi(k)^{-1}\exp(2\pi ikx).$$

Then, if $1 \leq p,q,r \leq +\infty$ with $1/p + 1/q = 1/r + 1$,

$$\begin{aligned}\left\{\int_{\mathbb{T}^d}\left|\int_{\mathbb{T}^d} f(x-y)\,d\mu(y)\right|^r dx\right\}^{1/r} &= \left\{\int_{\mathbb{T}^d}|B*A*f*\mu(x)|^r\,dx\right\}^{1/r} \\ &\leq \left\{\int_{\mathbb{T}^d}|A*f(x)|^p\,dx\right\}^{1/p}\left\{\int_{\mathbb{T}^d}|B*\mu(x)|^q\,dx\right\}^{1/q}.\end{aligned}$$

In the case of the sphere $\mathbb{S}^d = \left\{x \in \mathbb{R}^{d+1},\ |x| = 1\right\}$, let $\{Z_n(xy)\}$ be the system of zonal spherical harmonics polynomials and let

$$A(xy) = \sum_{n=0}^{+\infty}\psi(n)Z_n(xy), \quad B(xy) = \sum_{n=0}^{+\infty}\psi(n)^{-1}Z_n(xy).$$

Then, if $1 \leq p,q \leq +\infty$ with $1/p + 1/q = 1$,

$$\begin{aligned}&\left|\int_{\mathbb{S}^d} f(x)\,d\mu(x)\right| \\ &\leq \left\{\int_{\mathbb{S}^d}\left|\int_{\mathbb{S}^d} A(xy)f(y)dy\right|^p dx\right\}^{1/p}\left\{\int_{\mathbb{S}^d}\left|\int_{\mathbb{S}^d} B(xy)d\mu(y)\right|^q dx\right\}^{1/q}.\end{aligned}$$

Both results on the torus and the sphere follow from Young inequality for convolutions.

Remark 3.5. A result related to Theorem 2.1 is the following. Identify $\mathbb{T}^d$ with the unit cube $\{0 \leq x_j < 1\}$ and denote by $\chi_{P(y)}(x)$ the characteristic function of the parallelepiped $P(y) = \{0 \leq x_j < y_j\}$. Then define

$$B(x) = \int_{\mathbb{T}^d} \chi_{P(y)}(x)dy - 2^{-d} = \prod_{j=1}^{d} (1 - x_j) - 2^{-d}$$
$$= \sum_{k \in \mathbb{Z}^d - \{0\}} \left(\left(\prod_{k_j = 0} 2 \right) \left(\prod_{k_j \neq 0} 2\pi i k_j \right) \right)^{-1} \exp(2\pi i k x).$$

Also, define the differential integral operator

$$A * f(x) = \sum_{k \neq 0} \left(\prod_{k_j = 0} 2 \right) \left(\prod_{k_j \neq 0} 2\pi i k_j \right) \widehat{f}(k) \exp(2\pi i k x)$$
$$= 2^{d-1} \sum_{1 \leq j \leq d} \int_{\mathbb{T}^{d-1}} \frac{\partial}{\partial x_j} f(x) \prod_{i \neq j} dx_i + 2^{d-2} \sum_{1 \leq i \neq j \leq d} \int_{\mathbb{T}^{d-2}} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \prod_{h \neq i,j} dx_h$$
$$... + \frac{\partial^d}{\partial x_1 ... \partial x_d} f(x).$$

Observe that, as in Theorem 2.1, the Fourier coefficients of the distribution $A(x)$ and of the function $B(x)$ are one inverse to the other, however here the Fourier coefficients are function of the lattice points $2\pi i k$, and not of the eigenvalues $4\pi^2 |k|^2$. If $d\nu(x) = N^{-1} \sum_{j=1}^{N} \delta_{z_j}(x)$, and if $1 \leq p, q, r \leq +\infty$ with $1/p + 1/q = 1/r + 1$, then

$$\left\{ \int_{\mathbb{T}^d} \left| N^{-1} \sum_{j=1}^{N} f(x - z_j) - \int_{\mathbb{T}^d} f(y)dy \right|^r dx \right\}^{1/r}$$
$$\leq \left\{ \int_{\mathbb{T}^d} |A * f(x)|^p dx \right\}^{1/p} \left\{ \int_{\mathbb{T}^d} |B * \nu(x)|^q dx \right\}^{1/q}.$$

The norm of $A * f(x)$ is dominated by an analogue of the Hardy Krause variation,

$$\left\{ \int_{\mathbb{T}^d} |A * f(x)|^p dx \right\}^{1/p}$$
$$\leq 2^{d-1} \sum_{1 \leq j \leq d} \left\{ \int_{\mathbb{T}} \left| \int_{\mathbb{T}^{d-1}} \frac{\partial}{\partial x_j} f(x) \prod_{i = j} dx_i \right|^p dx_j \right\}^{1/p}$$
$$+ 2^{d-2} \sum_{1 \leq i \neq j \leq d} \left\{ \int_{\mathbb{T}^2} \left| \int_{\mathbb{T}^{d-2}} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \prod_{h \neq i,j} dx_h \right|^p dx_i dx_j \right\}^{1/p}$$
$$... + \left\{ \int_{\mathbb{T}^d} \left| \frac{\partial^d}{\partial x_1 ... \partial x_d} f(x) \right|^p dx \right\}^{1/p}.$$

The norm of $B * \nu(x)$ is dominated by the discrepancy of the points $\{z_j\}_{j=1}^N$ with respect to the family of boxes $P(y)$,

$$\left\{ \int_{\mathbb{T}^d} |B * \nu(x)|^q \, dx \right\}^{1/q}$$
$$\leq \int_{\mathbb{T}^d} \left\{ \int_{\mathbb{T}^d} \left| N^{-1} \sum_{j=1}^{N} \chi_{P(y)} (z_j + x) - \prod_{j=1}^{d} y_j \right|^q dx \right\}^{1/q} dy.$$

In particular, the case $p = 1$ and $q = +\infty$ is an analogue of the Koksma Hlawka inequality. See [20]. A generalization of this classical inequality is contained in [6].

Remark 3.6. By Lemma 2.6 (1), the Bessel kernel $B^\alpha(x, y)$ with $\alpha > 0$ is a superposition of heat kernels $W(t, x, y)$. Indeed, it is possible to state an analogue of Corollary 2.4 in terms of the heat kernel, without explicit mention of Bessel potentials: If $\{z_j\}_{j=1}^N$ is a sequence of points in $\mathcal{M}$, if $\{\omega_j\}_{j=1}^N$ are positive weights with $\sum_j \omega_j = 1$, and if $f(x)$ is a function in $\mathbb{W}^{\alpha,p}(\mathcal{M})$ with $\alpha > d/2$, then

$$\left| \sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right|$$
$$\leq \left\{ \Gamma(\alpha)^{-1} \int_0^{+\infty} \left| \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j W(t, z_i, z_j) - 1 \right| t^{\alpha-1} \exp(-t) \, dt \right\}^{1/2} \|f\|_{\mathbb{W}^{\alpha,2}} .$$

This suggests the following heuristic interpretation: Mathematically, a set of points on a manifold is well-distributed if the associated Riemann sums are close to the integrals. Physically, a set of points is well-distributed if the heat, initially concentrated on them, in a short time diffuses uniformly across the manifold.

Remark 3.7. In order to minimize the errors in the numerical integration in Corollary 2.4 (3), one has to minimize the energies

$$\int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) \, d\nu(x) d\nu(y), \quad \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j B^{2\alpha}(z_i, z_j) .$$

These are analogous to the energy integrals in potential theory

$$\int_{\mathcal{M}} \int_{\mathcal{M}} |x - y|^{-\varepsilon} \, d\nu(x) d\nu(y).$$

See [14]. When $d < \alpha < d + 1$ the kernel $B^{2\alpha}(x, y)$ is positive and bounded, with a maximum at $x = y$ and a spike $A - B|x - y|^{2\alpha - d}$ when $x \to y$. In particular, the gradient at $x = y$ is infinite. This implies that in order to minimize the discrete energy $\sum_{i,j} \omega_i \omega_j B^{2\alpha}(z_i, z_j)$ the points $\{z_j\}$ have to be well separated. This suggests the following heuristic interpretation: Mathematically, a set of points on a manifold is well-distributed if the energy is minimal. Physically, a set of points, free to move and repelling each other according to some law, is well-distributed when they reach an equilibrium.

Remark 3.8. It can be proved that if $2\alpha > d + 2$ then

$$\left|B^{2\alpha}(x,x) - B^{2\alpha}(x,y)\right| \leq c\,|x-y|^2 .$$

This estimate in the proof of Theorem 2.7 yields that for most choices of sampling points $z_j \in U_j$,

$$\left|\sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx\right| \leq c \max_{1\leq j\leq N} \left\{\text{diameter}(U_j)^{d/2+1}\right\} \|f\|_{\mathbb{W}^{\alpha,2}(\mathcal{M})} .$$

The same result holds if $2\alpha = d + 2$, with a logarithmic transgression. Observe that these estimates hold for most choices of sampling points, but not for all choices. Indeed, if the manifold $\mathcal{M}$ is decomposed in disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup \ldots \cup U_N$ with measure $aN^{-1} \leq |U_j| = \omega_j \leq bN^{-1}$ and $\text{diameter}(U_j) \leq cN^{-1/d}$, if $f(x)$ is a smooth non constant function, and if the points $z_j \in U_j$ are the maxima of $f(x)$ in U_j, then $\sum_{j=1}^{N} \omega_j f(z_j)$ is an upper sum of the integral $\int_{\mathcal{M}} f(x)dx$ and

$$\sum_{j=1}^{N} \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \geq cN^{-1/d}.$$

Remark 3.9. Theorem 3.2 gives an estimate of the accuracy in a quadrature rule in terms of the discrepancy of a measure with respect to level sets of the Bessel kernel. The following argument shows that when the manifold is a sphere, or a rank one compact symmetric space, then the level sets of the heat kernel $\{W(t,x,y) > s\}$, and hence of the Bessel kernels $\{B^{\alpha}(x,y) \leq t\}$, are geodesic balls $\{|x-y| \leq r\}$. The Laplace operator on the sphere $\mathbb{S}^d$ with respect to a system of polar coordinates $x = (\vartheta, \sigma)$, with $0 \leq \vartheta \leq \pi$ the colatitude with respect to a given pole and $\sigma \in \mathbb{S}^{d-1}$ the longitude, is

$$\Delta_x = \Delta_{(\vartheta,\sigma)} = -\sin^{1-d}(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin^{d-1}(\vartheta)\frac{\partial}{\partial\vartheta}\right) + \Delta_\sigma.$$

Let $u(t,x)$ be the solution of the Cauchy problem for the heat equation

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,x) = -\Delta_x u(t,x), \\ u(0,x) = f(x). \end{cases}$$

If $f(x)$ depends only on the colatitude ϑ, if it is an even function decreasing in $0 < \vartheta < \pi$, then also $u(t,x)$ depends only on the colatitude and it is an even function decreasing in $0 < \vartheta < \pi$. In order to prove this, set $u(t,x) = U(t,\vartheta)$, $f(x) = F(\vartheta)$, and

$\sin^{d-1}(\vartheta)\,\partial U(t,\vartheta)/\partial\vartheta = V(t,\vartheta)$. Then

$$\begin{cases} \dfrac{\partial}{\partial\vartheta}\dfrac{\partial}{\partial t}U(t,\vartheta) = \dfrac{\partial}{\partial\vartheta}\left\{\sin^{1-d}(\vartheta)\,\dfrac{\partial}{\partial\vartheta}\left(\sin^{d-1}(\vartheta)\,\dfrac{\partial}{\partial\vartheta}U(t,\vartheta)\right)\right\}, \\ \dfrac{\partial}{\partial\vartheta}U(0,\vartheta) = \dfrac{\partial}{\partial\vartheta}F(\vartheta), \end{cases}$$

$$\begin{cases} \dfrac{\partial}{\partial t}V(t,\vartheta) = \dfrac{\partial^2}{\partial\vartheta^2}V(t,\vartheta) + (1-d)\,\dfrac{\cos(\vartheta)}{\sin(\vartheta)}\dfrac{\partial}{\partial\vartheta}V(t,\vartheta), \\ V(0,\vartheta) = \sin^{d-1}(\vartheta)\,\dfrac{\partial}{\partial\vartheta}F(\vartheta), \\ V(t,0) = V(t,\pi) = 0. \end{cases}$$

If $F(\vartheta)$ is decreasing in $0<\vartheta<\pi$, then $V(0,\vartheta)\le 0$ and, by the maximum principle, $V(t,\vartheta)\le 0$, hence $U(t,\vartheta)$ is decreasing in $0<\vartheta<\pi$. In particular, by considering a sequence of initial data $\{f_n(x)\}$ which depend only on the colatitude ϑ, even and decreasing in $0<\vartheta<\pi$, and which converge to the Dirac $\delta(x)$, one proves that the heat kernel $W(t,\cos(\vartheta))$ is decreasing in $0<\vartheta<\pi$. Since Bessel kernels are superposition of heat kernels, they are also superposition of spherical caps.

Remark 3.10. In [3] and [22] the discrepancy of orbits of discrete subgroups of rotations of a sphere are studied. Let $\mathcal{G}$ be a compact Lie group, $\mathcal{K}$ a closed subgroup, $\mathcal{M}=\mathcal{G}/\mathcal{K}$ a homogeneous space of dimension d. Also, let $\mathcal{H}$ be a finitely generated free subgroup in $\mathcal{G}$ and assume that the action of $\mathcal{H}$ on $\mathcal{M}$ is free. Given a positive integer n, let $\{\sigma_j\}_{j=1}^N$ be an ordering of the elements in $\mathcal{H}$ with length at most n and for every function $f(x)$ on $\mathcal{M}$, define

$$Tf(x) = N^{-1}\sum_{j=1}^N f(\sigma_j x).$$

This operator is self-adjoint and it has eigenvalues and eigenfunctions in $\mathbb{L}^2(\mathcal{M})$. Moreover, since the operators T and Δ commute, they have a common orthonormal system of eigenfunctions, $\Delta\varphi_\lambda(x)=\lambda^2\varphi_\lambda(x)$ and $T\varphi_\lambda(x)=T(\lambda)\varphi_\lambda(x)$. All eigenvalues of T have modulus at most 1 and indeed 1 is an eigenvalue and the constants are eigenfunctions. Assume that all non constant eigenfunctions have eigenvalues much smaller than 1. Then, if $\alpha>d/2$,

$$\left|N^{-1}\sum_{j=1}^N f(\sigma_j x) - \int_{\mathcal{M}} f(x)dx\right| = \left|\sum_{\lambda\ne 0} T(\lambda)\mathcal{F}f(\lambda)\varphi_\lambda(x)\right|$$

$$\le \left\{\sup_{\lambda\ne 0}\{|T(\lambda)|\}\right\}\left\{\sum_\lambda \left(1+\lambda^2\right)^\alpha |\mathcal{F}f(\lambda)|^2\right\}^{1/2}\left\{\sum_\lambda \left(1+\lambda^2\right)^{-\alpha}|\varphi_\lambda(x)|^2\right\}^{1/2}$$

$$\le c\left\{\sup_{\lambda\ne 0}\{|T(\lambda)|\}\right\}\left\{\int_{\mathcal{M}}\left|(I+\Delta)^{\alpha/2}f(x)\right|^2dx\right\}^{1/2}.$$

The absolute convergence of the above series is consequence of the Sobolev's imbeddings, or the Weyl's estimates for eigenfunctions. In particular, when $\mathcal{M} = SO(3)/SO(2)$ is the two dimensional sphere and $\mathcal{H}$ is the free group generated by rotations of angles $\arccos(-3/5)$ around orthogonal axes, it has been proved in [22] that the eigenvalues of the operator T satisfy the Ramanujan bounds

$$\sup_{\lambda \neq 0} \{|T(\lambda)|\} \leq cN^{-1/2} \log(N).$$

Hence, for the sphere,

$$\left| N^{-1} \sum_{j=1}^{N} f(\sigma_j x) - \int_{\mathcal{M}} f(x) dx \right| \\ \leq cN^{-1/2} \log(N) \left\{ \int_{\mathcal{M}} \left| (I + \Delta)^{\alpha/2} f(x) \right|^2 dx \right\}^{1/2}.$$

All of this is essentially contained in [22]. Although this bound $N^{-1/2} \log(N)$ is worse than the bound $N^{-\alpha/2}$ in Corollary 2.13, the matrices $\{\sigma_j\}$ have rational entries and the sampling points $\{\sigma_j x\}$ are completely explicit.

References

[1] Amstler, C., Zinterhof, P.: Uniform distribution, discrepancy, and reproducing kernel Hilbert spaces. J. Complexity 17 (2001), 497-515.

[2] Andrievskii, V. V., Blatt, H. P., Götz, M.: Discrepancy estimates on the sphere, Monatsh. Math. 128 (1999), 179-188.

[3] Arnold, V. I., Krylov, A. L.: Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain. Dokl. Akad. Nauk. SSSR 148 (1963), 9-12.

[4] Beck, J., Chen, W. W. L.: Irregularities of distribution. Cambridge University Press (1987).

[5] Brandolini, L., Colzani, L.: Decay of Fourier transforms and summability of eigenfunction expansions. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 29 (2000), 611-638.

[6] Brandolini, L., Colzani, L., Gigante, G., Travaglini, G.: On Koksma Hlawka inequality. In preparation.

[7] Brauchart, J. S., Hesse, K.: Numerical integration over spheres of arbitrary dimension. Constr. Approx. 25 (2007), 41-71.

[8] Brown, G., Feng, D., Sheng, S. Y.: Kolmogorov width of classes of smooth functions on the sphere S^{d-1}. J. Complexity 18 (2002), 1001-1023.

[9] Chavel, I.: Eigenvalues in Riemannian geometry. Academic Press (1984).

[10] Choirat, C., Seri, R.: Numerical properties of generalized discrepancies on spheres of arbitrary dimension. In preparation.

[11] Colzani, L., Gigante, G., Travaglini, G.: Trigonometric approximation and a general form of the Erdös Turan inequality. Trans. Amer. Math. Soc. 363 (2011), 1101-1123.

[12] Cui, J., Freeden, W.: Equidistribution on the sphere. SIAM J. Sci. Comput. 18 (1997), 595-609.

[13] Damelin, S. B., Levesley, J., Ragozin D. L., Sun, X.: Energies, group-invariant kernels and numerical integration on compact manifolds. J. Complexity 25 (2009), 152-162.

[14] Hardin, D. P., Saff, E. B.: Discretizing manifolds via minimum energy points. Notices Amer. Math. Soc. 51 (2004), 1186-1194.

[15] Hesse, K., Sloan, I. H.: Optimal lower bounds for cubature error on the sphere S^2. J. Complexity 21 (2005), 790-803.

[16] Hesse, K., Sloan, I. H.: Cubature over the sphere S^2 in Sobolev spaces of arbitrary order. J. Approx. Theory 141 (2006), 118-133.

[17] Hesse, K., Mhaskar, H. N., Sloan, I. H.: Quadrature in Besov spaces on the Euclidean sphere. J. Complexity 23 (2007), 528-552.

[18] Hörmander, L.: The analysis of linear partial differential operators, I, II, III, IV. Springer Verlag (1983-1985).

[19] Kuijlaars, A. B. J., Saff, E. B.: Distributing many points on a sphere. Math. Intelligencer 19 (1997), 5-11.

[20] Kuipers, L., Niederreiter, H.: Uniform distribution of sequences. Pure and Applied Mathematics, Wiley-Interscience (1974).

[21] Kushpel, A.: Optimal cubature formulas on compact homogeneous manifolds. J. Funct. Anal. 257 (2009), 1621-1629.

[22] Lubotzky, A., Phillips R., Sarnak, P.: Hecke operators and distributing points on the sphere I, II, Comm. Pure Appl. Math. 39 (1986), 149-186, 40 (1987), 401-420.

[23] Matoušek, J.: Geometric discrepancy. An illustrated guide. Springer Verlag (1999).

[24] Novak, E.: Deterministic and stochastic error bounds in numerical analysis. Lecture Notes Mathematics 1349, Springer Verlag (1988).

[25] Seymour, P. D., Zaslavsky, T.: Averaging sets: a generalization of mean values and spherical designs. Adv. in Math. 52 (1984), 213-240.

[26] Shapiro, H. S.: Topics in approximation theory. Lecture Notes Math. 187, Springer Verlag (1971).

[27] Taylor, M. E.: Pseudodifferential operators. Princeton University Press (1981).