J Fourier Anal Appl (2010) 16: 629–657
DOI 10.1007/s00041-010-9119-4

A Quadrature Formula for Diffusion Polynomials Corresponding to a Generalized Heat Kernel

F. Filbir · H.N. Mhaskar

Received: 20 October 2008 / Revised: 18 September 2009 / Published online: 12 February 2010

Abstract Let $\{\phi_k\}$ be an orthonormal system on a quasi-metric measure space $\mathbb{X}$, $\{\ell_k\}$ be a nondecreasing sequence of numbers with $\lim_{k\to\infty}\ell_k = \infty$. A diffusion polynomial of degree L is an element of the span of $\{\phi_k : \ell_k \le L\}$. The heat kernel is defined formally by $K_t(x,y) = \sum_{k=0}^\infty \exp(-\ell_k^2 t)\phi_k(x)\overline{\phi_k(y)}$. If T is a (differential) operator, and both K_t and $T_y K_t$ have Gaussian upper bounds, we prove the Bernstein inequality: for every p, $1 \le p \le \infty$ and diffusion polynomial P of degree L, $\|TP\|_p \le c_1 L^c \|P\|_p$. In particular, we are interested in the case when $\mathbb{X}$ is a Riemannian manifold, T is a derivative operator, and $p \ne 2$. In the case when $\mathbb{X}$ is a compact Riemannian manifold without boundary and the measure is finite, we use the Bernstein inequality to prove the existence of quadrature formulas exact for integrating diffusion polynomials, based on an *arbitrary* data. The degree of the diffusion polynomials for which this formula is exact depends upon the mesh norm of the data. The results are stated in greater generality. In particular, when T is the identity operator, we recover the earlier results of Maggioni and Mhaskar on the summability of certain diffusion polynomial valued operators.

Communicated by Karlheinz Gröechenig.

The research of F. Filbir was partially funded by Deutsche Forschungsgemeinschaft grant FI 883/3-1 and PO711/9-1.
The research of H.N. Mhaskar was supported, in part, by grant DMS-0605209 and its continuation DMS-0908037 from the National Science Foundation and grant W911NF-09-1-0465 from the US Army Research Office.

F. Filbir (✉)
Institute of Biomathematics and Biometry, Helmholtz Center Munich, 85764 Neuherberg, Germany
e-mail: filbir@helmholtz-muenchen.de

H.N. Mhaskar
Department of Mathematics, California State University, Los Angeles, CA 90032, USA
e-mail: hmhaska@calstatela.edu

Keywords Approximation on manifolds · Bernstein inequalities · Marcinkiewicz-Zygmund inequalities · Quadrature formulas

Mathematics Subject Classification (2000) 41A17 · 41A55 · 58J35 · 58J90

1 Introduction

Many practical applications, for example, document analysis [6], face recognition [17], semi-supervised learning [1, 2], image processing [10], and cataloguing of galaxies [11], involve a large amount of very high dimensional data. Typically, this data has a lower intrinsic dimensionality; for example, one may assume that it belongs to a low dimensional manifold in a high dimensional, ambient Euclidean space. The desire to take advantage of this low intrinsic dimensionality has recently prompted a great deal of research on diffusion geometry techniques. The special issue [5] of Applied and Computational Harmonic Analysis contains several papers that serve as a good introduction to this subject.

An essential ingredient in these works is a data-dependent *heat kernel* K_t on the manifold $\mathbb{X}$ in question, which can be defined formally by

$$K_t(x,y)=\sum_{k\ge 0}\exp(-\ell_k^2 t)\phi_k(x)\overline{\phi_k(y)},\quad t>0, x, y\in\mathbb{X},$$

where the *eigenfunctions* $\{\phi_k\}$ are an orthonormal basis for $L^2(\mathbb{X},\mu)$ for an appropriate measure μ, and ℓ_k's are nonnegative numbers; the eigenvalues of the (square root of the negative) *Laplacian*. A multiresolution analysis is then defined by Coifman and Maggioni [6] for a fixed $\epsilon>0$ by defining the increasing sequence of scaling spaces

$$\mathrm{span}\{\phi_k:\exp(-2^{-j}\ell_k^2)\ge\epsilon\}=\mathrm{span}\{\phi_k:\ell_k^2\le(2^j\log(1/\epsilon))\}.$$

The range of the operators generated by $K_{2^{-j}}$ being "close" to the space at level j, one may obtain an approximate projection of a function by applying these operators to the function. In turn, these operators can be computed using fast multipole techniques. The diffusion wavelets and wavelet packets can be obtained by applying Gram Schmidt procedure to the kernels $K_{2^{-j}}$. On a more theoretical side, Jones, Maggioni, and Schul [18] have recently proved that the heat kernel can be used to construct a local coordinate atlas on manifolds, preserving the order of magnitude of the distances between points within each chart.

Several applications, especially in the context of semi-supervised learning, signal processing, and pattern recognition can be viewed as problems of function approximation. For example, given a few digitized images of handwritten digits, one wishes to develop a model that will predict for any other image whether the corresponding digit is 0. Each image may be viewed as a point in a high dimensional space, and the target function is the characteristic function of the set of points corresponding to the digit 0. We observe in this context that even though $\mathcal{K}_t f\to f$ (uniformly if f is continuous) as $t\to 0$, where $\mathcal{K}_t$ is the heat operator defined by the kernel K_t, the rate of convergence provided by this simple minded approximation cannot be the optimal one for smooth functions, since the $\mathcal{K}_t\phi_k\ne\phi_k$ except when $\ell_k=0$. In [24, 27], we have developed a different multiscale analysis based on $\mathsf{span}\{\phi_k : \ell_k^2 \le 2^j\}$ as the scaling spaces. We have obtained a Littlewood–Paley expansion, valid for functions in all L^p spaces *including* $p = 1, \infty$. This expansion is in terms of a tight frame transform, which can be used to characterize different Besov spaces. Our tight frames can also be chosen to be highly localized. The critical conditions in our paper [24] to ensure the uniform boundedness of the approximation operators involved are the finite speed of wave propagation, and a "Christoffel function" bound of the following form, valid for all $x \in \mathbb{X}$:

$$\sum_{\ell_k \le j} |\phi_k(x)|^2 = \mathcal{O}(j^K), \quad \text{for some } K > 0. \tag{1.1}$$

In the sequel, we assume a fixed nondecreasing sequence $\{\ell_k\}_{k=0}^{\infty} \in \mathbb{R}$, with $\lim_{k\to\infty} \ell_k = \infty$. An element of the space $\Pi_L := \mathsf{span}\,\{\phi_k : \ell_k \le L\}$ will be called a diffusion polynomial of degree at most L. In order to implement our approximation operators based on the available data, without losing their approximation power, one needs to use carefully chosen quadrature formulas, exact for integrating products of diffusion polynomials. We note that the existence of quadrature formulas exact for integrating diffusion polynomials in Π_L, even with positive weights, is well known if one is allowed to choose the quadrature nodes judiciously [33, Example 2.5.8]. However, in most applications, one has no control on the choice of the quadrature nodes. Such data, where we have no control on the choice of the sites where the data is collected, will be referred to as scattered data. It is demonstrated in [23] that even in the simplest case when $\mathbb{X}$ is the unit circle of the complex plane, one cannot simply take a known formula based on judiciously chosen sites, such as the Fast Fourier Transform, and use it with finding nearest neighbors from the given scattered data, without losing the optimal approximation power. One major goal of this paper is to prove the existence of quadrature formulas to integrate diffusion polynomials of high order, based on scattered data.

In order to obtain the necessary quadrature formulas, we need to prove certain Bernstein inequalities, estimating the L^1 norm of the derivatives of diffusion polynomials in terms of the L^1 norms of these polynomials themselves. Such inequalities are well known in the case of L^2, and are recently investigated [32] in a more general context. As far as we are aware, these inequalities are not known in general for the derivatives in the L^p norm. In the context of trigonometric polynomials, a simple idea is to use a reproducing kernel, and estimate the L^1 norm of the derivative of this kernel. We wish to adopt this approach in the present context. Since the derivatives of ϕ_k are not necessarily diffusion polynomials, we need to drop the assumptions on orthogonality of the system and consider generalized heat kernels of the form $\sum_{k=0}^{\infty} \exp(-\ell_k^2 t)\phi_k(x)\overline{\psi_k(y)}$, where ϕ_k, ψ_k are no longer required to be orthogonal systems. Another motivation for considering such a kernel is the following. The spectral method for solving a linear differential equation $Tu = f$ on a manifold consists of solving a system of equations

$$T\left(\sum_{k=0}^{n} \langle u, \phi_k\rangle \phi_k\right) = \sum_{k=0}^{n} \langle u, \phi_k\rangle T\phi_k = \sum_{k=0}^{n} \langle f, \phi_k\rangle \phi_k,$$

for the unknown coefficients $\langle u, \phi_k\rangle$. In the analysis of the middle term above, the kernel has the form $\sum_{k=0}^{n} \phi_k(x)\overline{\psi_k(y)}$, where $\psi_k = T\phi_k$. Thus, the systems ϕ_k and ψ_k

are not the same and not orthogonal. Other examples of interest include bi-orthogonal frame expansions.

In Sect. 2, our main goal is to extend the localization and summability result [24, Theorem 4.1] to the case of kernels of the form $\sum_{k=0}^{\infty} H(\ell_k/L)\phi_k(x)\overline{\psi_k(y)}$ for suitable filters H (Theorem 2.1). In light of the importance of the heat kernel in this theory, we will formulate the conditions we need for this purpose only in terms of the small time properties of the (generalized) heat kernel. Our conditions on the heat kernel are satisfied in a number of situations; for example, weighted heat kernels, heat kernels corresponding to a wide class of elliptic operators of second order on smooth manifolds with bounded geometry, and of course, kernels corresponding to the Laplace–Beltrami operators on smooth, compact manifolds (without boundary), together with the gradients of these kernels. This theory will be applied to prove in Theorem 2.2 the Bernstein inequalities in the contexts when the conditions on the heat kernel and its gradient are satisfied. Restricting further to compact Riemannian manifolds with a probability measure, we will prove in Theorem 3.1 the existence of the quadrature formulas. The proofs of all the new results of this paper are given in Sect. 4. For the convenience of the reader, we include the definitions of some of the terminology regarding Riemannian manifolds in this paper in an Appendix.

2 Main Results

Let $(\mathbb{X}, \mu, \rho)$ be a quasi-metric measure space, with a quasi-metric ρ and a sigma-finite Borel measure μ. By the term quasi-metric, we mean that $\rho : \mathbb{X} \times \mathbb{X} \to [0, \infty)$ is a symmetric function, $\rho(x, y) = 0$ if and only if $x = y$, and the following inequality is satisfied in place of the triangle inequality for some $\kappa > 0$:

$$\rho(x, y) \le \kappa\{\rho(x, z) + \rho(z, y)\}, \quad x, y, z \in \mathbb{X}. \tag{2.1}$$

For $x \in \mathbb{X}$, $r > 0$, let

$$B(x, r) := \{y \in \mathbb{X} : \rho(x, y) \le r\}, \qquad \Delta(x, r) := \mathbb{X} \setminus B(x, r).$$

We assume that there exists a constant $\kappa_1 > 0$ such that

$$\mu(B(x, r)) \le \kappa_1 r^{\alpha}, \quad x \in \mathbb{X}, r > 0. \tag{2.2}$$

If $B \subseteq \mathbb{X}$ is μ-measurable, and $f : B \to \mathbb{C}$ is a μ-measurable function, we will write

$$\|f\|_{p,B} := \begin{cases} \{\int_B |f(x)|^p d\mu(x)\}^{1/p}, & \text{if } 1 \le p < \infty, \\ \mu - \text{ess sup}_{x \in B} |f(x)|, & \text{if } p = \infty. \end{cases}$$

The class of all f with $\|f\|_{p,B} < \infty$ will be denoted by $L^p(B)$, with the usual convention of considering two functions to be equal if they are equal μ-almost everywhere. If $B = \mathbb{X}$, we will omit its mention from the notations. The inner product of L^2 will be denoted by $\langle \circ, \circ \rangle$. We denote the L^p closure of $\Pi_\infty := \bigcup_{L \ge 0} \Pi_L$ by X^p.

We equip the space $L^1 \cap X^\infty$ with the norm $\| \circ \|_1 + \| \circ \|_\infty$. It is elementary to check that $L^1 \cap X^\infty \subset L^p$ for every p, $1 \le p \le \infty$, the inclusion being a norm embedding. A system $\{\phi_k\} \subset L^2$ will be called a *Bessel system* if there exists a dense subset $\mathcal{D} = \mathcal{D}(\{\phi_k\})$ of $L^1 \cap X^\infty$ (with respect to the norm of this space), such that (i) for any $\epsilon > 0$, ball of the form $B(x,r)$, and $f \in L^1 \cap X^\infty$ supported on $B(x,r)$, there exists $g \in \mathcal{D}$ such that the support of g is contained in $B(x,2r)$ and $\|f - g\|_1 \le \epsilon$, and (ii)

$$\sum_{k=0}^{\infty} |\langle f, \phi_k\rangle|^2 \le \mathcal{N}(f) < \infty, \quad f \in \mathcal{D}, \tag{2.3}$$

where $\mathcal{N}(f)$ is a positive number dependent on f, $\langle \circ, \circ \rangle$ and $\{\phi_k\}$. Obviously, any orthonormal system on $\mathbb{X}$ is a Bessel system with $\mathcal{D} = L^1 \cap X^\infty$. Another interesting example is the following. Let $\mathbb{X}$ be a Riemannian manifold, $\{\phi_k\}$ be the eigenfunctions of the Laplace–Beltrami operator on $\mathbb{X}$, F be a conservative vector field on $\mathbb{X}$, and $\psi_k = F\phi_k$, $k = 0, 1, \ldots$. For the space $\mathcal{D}$ we choose the class of all compactly supported, infinitely differentiable functions on $\mathbb{X}$. Then an integration by parts argument shows that (2.3) is satisfied for $\{\psi_k\}$ as well, with $\mathcal{N}(f)$ being the L^2 norm of the corresponding derivative of f.

Let $\{\phi_k\}$, $\{\psi_k\}$ be Bessel systems, each of whose members is assumed to be continuous, and in $L^1 \cap X^\infty$. Let $\{\ell_k\}$ be an nondecreasing sequence of nonnegative numbers, with $\ell_k \uparrow \infty$ as $k \to \infty$. An element of $\mathsf{span}\,\{\phi_k : \ell_k \le L\}$ will be called a *diffusion polynomial* (of degree at most L). We note that this terminology distinguishes ϕ_k's from ψ_k's, which may not be diffusion polynomials. For $t > 0$, the *heat kernel* is defined formally by

$$K_t(\{\phi_k\}, \{\psi_k\}; x, y) = \sum_{k=0}^{\infty} \exp(-\ell_k^2 t)\phi_k(x)\overline{\psi_k(y)}. \tag{2.4}$$

We assume that there exist constants $\kappa_2, \ldots, \kappa_5, A_1, A_2, A_3 > 0$ such that for all $x, y \in \mathbb{X}$ and $t \in (0, 1]$,

$$K_t(\{\phi_k\}, \{\phi_k\}; x, x) \le \kappa_2 t^{-A_1/2}, \qquad K_t(\{\psi_k\}, \{\psi_k\}; x, x) \le \kappa_3 t^{-A_2/2}, \tag{2.5}$$

$$|K_t(\{\phi_k\}, \{\psi_k\}; x, y)| \le \kappa_4 t^{-A_3} \exp(-\kappa_5 \rho(x,y)^2/t). \tag{2.6}$$

In the remainder of this paper, the symbols $c, c_1, \ldots$ will denote generic positive constants depending only on the fixed parameters in the discussion, such as the space $\mathbb{X}$ and related quantities like ρ, μ, $\kappa, \kappa_1, \ldots, \kappa_5$, A_1, A_2, and the norms to be introduced below. These constants do not depend upon the systems $\{\phi_k\}$, $\{\psi_k\}$ themselves except through the constants $\kappa, \kappa_1, \ldots, \kappa_5$, A_1 and A_2. They are also independent of the functional $\mathcal{N}$ in (2.3). *Their value may be different at different occurrences, even within a single formula. The notation $A \sim B$ will mean $c_1 A \le B \le c_2 A$.*

In view of the Schwarz inequality, an obvious consequence of (2.5) is that the heat kernel is well defined as a function on $(0,1] \times \mathbb{X} \times \mathbb{X}$. We will see in Proposition 4.1 below that the first and second estimate in (2.5) are equivalent to the first and second estimate respectively in

$$\sum_{\ell_k \le u} |\phi_k(x)|^2 \le c u^{A_1}, \quad \sum_{\ell_k \le u} |\psi_k(x)|^2 \le c u^{A_2}, \quad u \ge 1. \tag{2.7}$$

Sikora [35] has proved (cf. Theorem 4.1 below) that the condition (2.6) is equivalent to the finite speed of wave propagation. To explain this notion, let f_1 (respectively, f_2) be a function for which (2.3) holds with $\{\phi_k\}$ (respectively, $\{\psi_k\}$). Then for $t \in \mathbb{R}$, the *wave kernel* defined by

$$W(t, f_1, f_2) := W(\{\phi_k\}, \{\psi_k\}; t, f_1, f_2) := \sum_{k=0}^{\infty} \cos(\ell_k t)\langle f_1, \phi_k\rangle \overline{\langle f_2, \psi_k\rangle} \tag{2.8}$$

is well defined. The finite speed of wave propagation means that $W(t, f_1, f_2) = 0$ if $|t|$ does not exceed a constant multiple (the speed) of the distance between the supports of f_1 and f_2. In particular, the conditions of [24, Theorem 4.1] related to the eigenfunctions and the wave kernel are equivalent to the Gaussian bound (with α as in (2.2))

$$|K_t(\{\phi_k\}, \{\phi_k\}; x, y)| \le c_1 t^{-\alpha/2} \exp(-c_3\rho(x, y)^2/t), \quad x, y \in \mathbb{X}, t \in (0, 1]. \tag{2.9}$$

Thus, Theorem 2.1 below generalizes [24, Theorem 4.1], at the same time clarifying some of the assumptions made in that theorem. A further motivation for stating the theorem in this generality stem from the following considerations. In [37], Bin Xu has proved that in the case when $\mathbb{X}$ is a smooth, closed Riemannian manifold (i.e., compact Riemannian manifold without boundary) with dimension n, ∂^m is a derivative of order m, each ϕ_k is the eigenfunction of the Laplace–Beltrami operator corresponding to the eigenvalue ℓ_k^2, and $\psi_k = \partial^m \phi_k$, then (2.7) holds with $A_1 = \alpha = n$, $A_2 = n + 2m$, and the constants depending only on m and not on the particular derivative. Since the finite speed of wave propagation holds in the case of $W(\{\phi_k\}, \{\phi_k\})$, so does it hold for the derivative kernel $W(\{\phi_k\}, \{\psi_k\})$. Kordyukov [20] has proved similar results for the heat kernels corresponding to a very general class of second order elliptic operators (in place of the Laplace–Beltrami operator) on a Riemannian manifold with "bounded geometry" (see [20] for definitions). Estimates on the heat kernel and its gradients are well understood in many other cases, including higher order partial differential operators on manifolds [4, 7, 12, 14], with many other references given in [15].

Theorem 2.1 *Let* $\{\phi_k\}, \{\psi_k\}$ *be Bessel systems*, (2.2), (2.5), *and* (2.6) *hold*, $K = (A_1 + A_2)/2$, $S > \max(K, \alpha)$ *be an integer*, $L > 0$. *Let* $H : \mathbb{R} \to \mathbb{R}$ *be an even function, supported on* $[-1, 1]$ *with continuous derivatives of order* S. *We normalize* H *so that*

$$\sum_{k=0}^{S} \max_{u \in \mathbb{R}} |H^{(k)}(u)| = 1.$$

Then for $x, y \in \mathbb{X}$,

$$\left| \sum_{k=0}^{\infty} H(\ell_k/L)\phi_k(x)\overline{\psi_k(y)} \right| \le c \frac{L^K}{\max(1, (L\rho(x, y))^S)}. \tag{2.10}$$

Consequently,

$$\sup_{x\in\mathbb{X}}\int_{\mathbb{X}}\left|\sum_{k=0}^{\infty}H(\ell_k/L)\phi_k(x)\overline{\psi_k(y)}\right|d\mu(y)\le cL^{K-\alpha},$$
$$\sup_{y\in\mathbb{X}}\int_{\mathbb{X}}\left|\sum_{k=0}^{\infty}H(\ell_k/L)\phi_k(x)\overline{\psi_k(y)}\right|d\mu(x)\le cL^{K-\alpha}. \tag{2.11}$$

Therefore, for any p, $1\le p\le\infty$, *and* $f\in L^p$,

$$\left\|\sum_k H(\ell_k/L)\langle f,\phi_k\rangle\psi_k\right\|_p\le cL^{K-\alpha}\|f\|_p,$$
$$\left\|\sum_k H(\ell_k/L)\langle f,\psi_k\rangle\phi_k\right\|_p\le cL^{K-\alpha}\|f\|_p. \tag{2.12}$$

As an application of the above theorem, we obtain the following Bernstein inequality (2.14) for diffusion polynomials. In particular, the following theorem is valid in the situations considered in the papers cited above, where Gaussian bounds on the gradients of the heat kernel are known.

Theorem 2.2 *Let* $\{\phi_k\}$ *be a complete orthonormal system in* L^2, T *be a linear operator defined for all diffusion polynomials, and* $\psi_k=T\phi_k$, *and the assumptions in Theorem 2.1 hold. Let* $L>0$, C_1, C_2 *be open subsets of* $\mathbb{X}$ *with the closure of* C_1 *being a compact subset of* C_2, *and* $d=\inf_{y\in\mathbb{X}\setminus C_2}\rho(y,C_1)$. *Then for any* $P\in\Pi_L$ *and any* $s>\max(K,\alpha)$, *we have*

$$\|TP\|_{p,C_1}\le cL^{K-\alpha}\{\|P\|_{p,C_2}+(\max(1,Ld))^{-S}\|P\|_p\}. \tag{2.13}$$

In particular,

$$\|TP\|_p\le cL^{K-\alpha}\|P\|_p. \tag{2.14}$$

As mentioned before, in many cases, we may choose T to be a derivative ∂^m on a manifold and $K=\alpha+m$. In these cases, the assumptions of Theorem 2.1 may be consolidated into (2.2), (2.9), and

$$|\partial^m K_t(\{\phi_k\},\{\phi_k\};x,y)|\le ct^{-\alpha/2-m}\exp(-c_3\rho(x,y)^2/t),\quad x,y\in\mathbb{X},\ t\in(0,1]. \tag{2.15}$$

The estimate (2.14) then translates into

$$\|\partial^m P\|_p\le cL^m\|P\|_p,\quad L\ge 1,\ P\in\Pi_L, \tag{2.16}$$

and (2.13) translates into

$$\|\partial^m P\|_{p,C_1}\le cL^m\{\|P\|_{p,C_2}+(\max(1,Ld))^{-S}\|P\|_p\},\quad L\ge 1,\ P\in\Pi_L. \tag{2.17}$$

In particular, the estimates (2.16) and (2.17) are valid when P is a spherical polynomial of degree at most L on a Euclidean sphere, as well as similar settings on the rotation group $SO(3)$, and two point homogeneous spaces (see [13, 28, 34]). As far as we are aware, these results are new in the case when $p < \infty$.

3 Quadrature Formulas

In this section, we assume that $\mathbb{X}$ is a compact, connected, Riemannian manifold (without boundary), and ρ is the geodesic distance on $\mathbb{X}$. Our objective is to prove the existence of quadrature formulas based on scattered data on $\mathbb{X}$, exact for diffusion polynomials of a high degree. This degree is estimated in terms of the mesh norm of the data set, while the stability of the computational procedures will depend upon the minimal separation. The material which we will present in this section can be seen as a unifying approach for existing examples like the Euclidean sphere, the rotation group etc. [13, 28, 34]. At the same time, it provides a way for dealing with similar problems in a more general setting.

Let $\mathcal{C} \subset \mathbb{X}$ be a finite set. The *mesh norm* $\delta_{\mathcal{C}}$ of $\mathcal{C}$ and the minimal separation $q_{\mathcal{C}}$ are defined by

$$\delta_{\mathcal{C}} = \sup_{x\in\mathbb{X}} \rho(x, \mathcal{C}), \;\; q_{\mathcal{C}} = \min_{x,y\in\mathcal{C}, x\neq y} \rho(x, y). \tag{3.1}$$

Since $\mathbb{X}$ is compact, for any $\delta > 0$, there is a finite set $\mathcal{C}$ with $\delta_{\mathcal{C}} \le \delta$. The condition (2.2) on the measure μ then implies that $\mu(\mathbb{X}) < \infty$. We will assume in this section that μ is normalized to be a probability measure. We will assume further that $\ell_0 = 0$, and $\phi_0(x) \equiv 1$ for every $x \in \mathbb{X}$. This ensures that

$$\int_{\mathbb{X}} K_t(\{\phi_k\}, \{\phi_k\}; x, y)d\mu(y) = 1, \quad x \in \mathbb{X}. \tag{3.2}$$

First we construct a reduced set $\tilde{\mathcal{C}}$ from the given set $\mathcal{C}$ which has the property $\delta_{\tilde{\mathcal{C}}} \le 2q_{\tilde{\mathcal{C}}}$. If $\mathcal{C} = \{x_1, \dots, x_M\}$, we let $\mathcal{C}_1 = \mathcal{C} \cap \Delta(x_1, \delta_{\mathcal{C}})$. By relabeling the set if necessary, we choose $x_2 \in \mathcal{C}_1$, and set $\mathcal{C}_2 = \mathcal{C}_1 \cap \Delta(x_2, \delta_{\mathcal{C}})$. Necessarily, $\rho(x_1, \mathcal{C}_2) \ge \delta_{\mathcal{C}}$ and $\rho(x_1, x_2) \ge \delta_{\mathcal{C}}$. Since $\mathcal{C}$ is finite, we may continue in this way at most M times to obtain a subset $\tilde{\mathcal{C}}$ of $\mathcal{C}$ which contains the points $x_1, x_2, \dots$ used for the construction. For $\tilde{\mathcal{C}}$ we have $q_{\tilde{\mathcal{C}}} \ge \delta_{\mathcal{C}}$, and moreover, for any $x \in \mathcal{C}$, there is $y \in \tilde{\mathcal{C}}$ with $\rho(x, y) \le \delta_{\mathcal{C}}$. It follows that $\delta_{\mathcal{C}} \le \delta_{\tilde{\mathcal{C}}} \le 2\delta_{\mathcal{C}} \le 2q_{\tilde{\mathcal{C}}}$. In the sequel, we will only work with the subset $\tilde{\mathcal{C}}$. Since the rest of the points in $\mathcal{C}$ are ignored in our analysis, we may rename this subset again as $\mathcal{C}$ and assume more generally that there exists $\kappa_6 > 0$ such that

$$\delta_{\mathcal{C}} \le \kappa_6 q_{\mathcal{C}}. \tag{3.3}$$

The constants in this section will depend also on κ_6.

Theorem 3.1 *Let $\mathbb{X}$ be a compact, connected, Riemannian manifold* (*without boundary*), *ρ be the geodesic distance on $\mathbb{X}$, μ be a probability measure on $\mathbb{X}$, and $\{\phi_k\}$ be an orthonormal system of continuous functions. We assume further that* (2.2), (2.9) *hold, and further that* (2.15) *holds with* $m = 1$ *for every differential operator of first order* ∂. *Let* $\mathcal{C}$ *be a finite subset of* $\mathbb{X}$ *satisfying* (3.3) *and* $\delta_{\mathcal{C}} \le 1/6$.

(a) *There exists* $c > 0$ *with the following property: for* $L \le c\delta_{\mathcal{C}}^{-1}$, *there exist numbers* w_x, $x \in \mathcal{C}$, *such that for each* $x \in \mathcal{C}$, $|w_x| \le c_2\mu(B(x,\delta_{\mathcal{C}})) \le c_3\delta_{\mathcal{C}}^{\alpha} \le c_4 q_{\mathcal{C}}^{\alpha}$, *and*

$$\int_{\mathbb{X}} P(y)d\mu(y) = \sum_{x\in\mathcal{C}} w_x P(x), \qquad P \in \Pi_L. \tag{3.4}$$

(b) *There exists* $c > 0$ *with the following property: for* $L \le c\delta_{\mathcal{C}}^{-1}$, *there exist positive numbers* $w_x^+ \ge c_1\mu(B(x,\delta_{\mathcal{C}}))$, $x \in \mathcal{C}$ *such that* (3.4) *holds with* w_x^+ *in place of* w_x.

Our proof of Theorem 3.1 follows arguments similar to those in [26, 28]. An essential step in this argument, which requires a proof very different from that in these papers, is the following Marcinkiewicz–Zygmund inequalities.

Theorem 3.2 *We assume the setup as in Theorem* 3.1. *Then there exists* $c > 0$ *such that for every* $\eta > 0$, $L \le c\eta\delta_{\mathcal{C}}^{-1}$, *and* $P \in \Pi_L$,

$$\begin{aligned}\sum_{x\in\mathcal{C}} \int_{B(x,\delta_{\mathcal{C}})} &|P(z) - P(x)|d\mu(z) \\ &\le \sum_{x\in\mathcal{C}} \mu(B(x,\delta_{\mathcal{C}})) \sup_{z\in B(x,\delta_{\mathcal{C}})} |P(z) - P(x)| \le \eta\|P\|_1.\end{aligned} \tag{3.5}$$

In particular,

$$c_1\|P\|_1 \le \sum_{x\in\mathcal{C}} \mu(B(x,\delta_{\mathcal{C}}))|P(x)| \le c_2\|P\|_1, \qquad P \in \Pi_L. \tag{3.6}$$

Moreover, if $P(x) \ge 0$ *for all* $x \in \mathcal{C}$, *then*

$$\int_{\mathbb{X}} P(z)d\mu(z) \ge c_3 \sum_{x\in\mathcal{C}} \mu(B(x,\delta_{\mathcal{C}}))P(x). \tag{3.7}$$

In [16], Grigor'yan has proved that (2.2), (3.2), and (2.9) together imply that

$$\mu(B(x,r)) \ge cr^{\alpha}, \qquad 0 < r \le 1, x \in \mathbb{X}. \tag{3.8}$$

Thus, we may replace the term $\mu(B(x,\delta_{\mathcal{C}}))$ in (3.6) by $\delta_{\mathcal{C}}^{\alpha}$ or $q_{\mathcal{C}}^{\alpha}$, of course, obtaining different constants. Also, using (2.2) and (3.8), we obtain that μ satisfies the homogeneity condition

$$\mu(B(x,R)) \le c(R/r)^{\alpha}\mu(B(x,r)), \qquad x \in \mathbb{X}, 0 < r \le 1, R > 0. \tag{3.9}$$

We observe an immediate consequence of Theorem 3.1. Since $|w_y| \le c_2\mu(B(y,\delta_{\mathcal{C}}))$, $y \in \mathcal{C}$, (3.6) implies that

$$\sum_{y\in\mathcal{C}} |w_y||P(y)| \le c_1\|P\|_1, \quad P \in \Pi_L. \tag{3.10}$$

Let

$$\Phi_L(x,y) := \Phi_L(H;x,y) := \sum_{k=0}^{\infty} H(\ell_k/L)\phi_k(x)\overline{\phi_k(y)}, \quad x,y \in \mathbb{X}, L>0. \tag{3.11}$$

Using $A_1 = A_2 = \alpha$ in the estimate (2.11) we obtain from (3.10) that

$$\sup_{x\in\mathbb{X}} \sum_{y\in\mathcal{C}} |w_y||\Phi_L(x,y)| \le \sup_{x\in\mathbb{X}} \int_{\mathbb{X}} |\Phi_L(x,y)|d\mu(y) \le c. \tag{3.12}$$

Using the Riesz–Thorin interpolation theorem, one can deduce from here that the discretized operator, defined for any $f : \mathbb{X} \to \mathbb{R}$ by

$$\sigma_L(\mathcal{C}; H, f, x) := \sum_{y\in\mathcal{C}} w_y f(y)\Phi_L(x,y),$$

satisfies

$$\|\sigma_L(\mathcal{C}; H, f)\|_p \le c\left\{\sum_{y\in\mathcal{C}} |w_y||f(y)|^p\right\}^{1/p}, \quad 1 \le p \le \infty. \tag{3.13}$$

In many known manifolds, such as the unit interval, the unit circle, the unit sphere, the rotation group $SO(3)$, etc., the products of elements of Π_L are in Π_{cL} for some c. In such cases, the estimate (3.13) leads to the fact that the operators $\sigma_L(\mathcal{C}; H, f)$ provide a near best approximation from Π_{cL}. In the context of the Euclidean sphere, it is demonstrated by several numerical examples in [23] that in the presence of singularities in the target function f, this procedure provides a substantially better approximation on large parts of the sphere than the usual procedures of least square approximation. The superiority of this method over traditional least squares in certain aspects was also observed in [24] in the context of semi-supervised learning of hand-written digits.

A simple way to find the weights w_x is to solve the least square problem of minimizing $\sum w_x^2$ with the constraints $\sum_{x\in\mathcal{C}} w_x\phi_k(x) = \int_{\mathbb{X}} \phi_k d\mu$, $k = 0,\dots,L$ [23, Formula (4.1)]. To obtain nonnegative weights, one may augment this problem with the appropriate inequality constraints. Alternately, one may obtain w_x's so as to minimize

$$\sum_{\ell_k\le L}\left(\sum_{x\in\mathcal{C}} w_x\phi_k(x) - \int_{\mathbb{X}} \phi_k d\mu\right)^2.$$

(See [13, 19] for efficient computational strategies in the case of the Euclidean sphere.) In view of the localization of the kernel Φ_L, a better strategy seems to be to solve the system of equations $\sum_{x\in\mathcal{C}} w_x \Phi_L(x,y) = \int_{\mathbb{X}} \Phi_L(z,y)d\mu(z)$, $y\in\mathcal{C}$. Theorem 3.1 shows that each of the systems of equations involved in all these approaches has a solution.

4 Proofs

In this section we prove all the new theorems in this paper. In Sect. 4.1, we discuss the equivalence of the conditions on the heat kernel with the conditions assumed in [24]. In Sect. 4.2, we prove Theorems 2.1 and 2.2. Finally, Theorem 3.1 is proved in Sect. 4.3.

4.1 The Heat Kernel

First, we investigate the condition (2.5). In [29], Minakshisundaram has used the Wiener–Ikehara Tauberian theorem [21] to prove that if the heat kernel satisfies an asymptotic of the form

$$K_t(\{\phi_k\},\{\phi_k\};x,x) = ct^{-A/2}\{1+c_1t+c_2t^2+\cdots\}, \tag{4.1}$$

then $\lim_{L\to\infty} L^{-A}\sum_{\ell_k\le L}\phi_k^2(x)$ is a constant, independent of x. Korevaar [22] has proved that the Tauberian theorem does not hold if only an inequality is known in place of (4.1) and an inequality is expected in return. The following proposition is a "poor man's alternative", sufficient for our purpose here.

Proposition 4.1 *Let $\{a_j\}$ be a sequence of nonnegative numbers, $C>0$. Then*

$$\sup_{u\ge 1} u^{-2C}\sum_{\ell_j\le u} a_j^2 \le 2e \sup_{t\in(0,1]} t^C\sum_{j=0}^\infty \exp(-\ell_j^2t)a_j^2 \le e(\Gamma(C+1)+2)\sup_{u\ge 1} u^{-2C}\sum_{\ell_j\le u} a_j^2. \tag{4.2}$$

Proof In this proof only, let $s(u)=\sum_{\ell_j\le u} a_j^2$, $u>1$, and $s(u)=0$ if $u\le 1$. Then

$$\begin{aligned}\sum_{j=0}^\infty \exp(-\ell_j^2t)a_j^2 &= \sum_{\ell_j\le 1}\exp(-\ell_j^2t)a_j^2 + \int_{1+}^\infty e^{-v^2t}ds(v)\\ &= \sum_{\ell_j\le 1}\exp(-\ell_j^2t)a_j^2 + \int_0^\infty e^{-v^2t}ds(v).\end{aligned} \tag{4.3}$$

First, in this proof only, let $L\in(0,\infty)$ be chosen so that $\sup_{t\in(0,1]} t^C \times \sum_{j=0}^\infty \exp(-\ell_j^2t)a_j^2 < L < \infty$. An integration by parts shows that for any $u>0$,

$$Lt^{-C} \ge \int_0^u e^{-v^2t}ds(v) = s(u)e^{-u^2t} + 2t\int_0^u ve^{-v^2t}s(v)dv \ge s(u)e^{-u^2t}. \tag{4.4}$$

Using this estimate for $t = 1/u^2$, we obtain $s(u) \le Leu^{2C}$; i.e.,

$$\sum_{1<\ell_j \le u} a_j^2 \le Leu^{2C}, \quad u > 1. \tag{4.5}$$

Since for $t \in (0, 1]$,

$$\sum_{\ell_j \le 1} a_j^2 \le e^t \sum_{\ell_j \le 1} \exp(-\ell_j^2 t) a_j^2 \le Le,$$

the estimate (4.5) implies that $\sup_{u\ge 1} u^{-2C} \sum_{\ell_j \le u} a_j^2 \le 2eL$. We have thus proved the first inequality in (4.2).

To prove the second inequality in (4.2), let in this proof only, $B > 0$ be chosen so that $\sup_{u\ge 1} u^{-2C} \sum_{\ell_j \le u} a_j^2 < B < \infty$. In particular, $s(u) \le Bu^{2C}$ if $u \ge 1$. The fact that $s(u) = 0$ for $u \le 1$ implies that for any $u \ge 1$,

$$\begin{aligned}
\int_0^u e^{-v^2 t} ds(v) &= s(u)e^{-u^2 t} + 2t \int_0^u v e^{-v^2 t} s(v) dv \\
&= s(u)e^{-u^2 t} + 2t \int_1^u v e^{-v^2 t} s(v) dv \\
&\le Bu^{2C} e^{-u^2 t} + Bt \int_0^\infty v^{2C+1} e^{-v^2 t} dv \\
&= Bu^{2C} e^{-u^2 t} + \frac{B\Gamma(C+1)}{2} t^{-C}.
\end{aligned}$$

Letting $u \to \infty$, we obtain

$$\int_0^\infty e^{-v^2 t} ds(v) \le \frac{B\Gamma(C+1)}{2} t^{-C}, \quad t \in (0, 1]. \tag{4.6}$$

Since

$$\sum_{\ell_j \le 1} \exp(-\ell_j^2 t) a_j^2 \le \sum_{\ell_j \le 1} a_j^2 = s(1) \le B \le Bt^{-C}, \quad t \in (0, 1],$$

we obtain the second inequality in (4.2) from (4.6) and (4.3). □

Next, we turn our attention to the equivalence between (2.6) and the finite speed of wave propagation. The following Theorem 4.1 was proved by Sikora in [35]. We will present a simplified version of the statement and the proof. While Sikora's formulation of the statement is in terms of certain operators in a very general setting, the formulation we have chosen here is merely a property of complex numbers. For the proof, Sikora uses the Phragmen–Lindelöff theorem, and also refers to a particular version of the Paley–Wiener theorem. Our proof here is self-contained, and makes use only of the identity theorem of complex analysis.

For an integrable function $f : \mathbb{R} \to \mathbb{R}$, its Fourier transform is defined by

$$\hat{f}(t) = \int_{\mathbb{R}} f(u) e^{-iut} du, \quad t \in \mathbb{R}, \tag{4.7}$$

and we recall that if both f and $\hat{f}$ are integrable then the Fourier inversion formula holds:

$$f(u) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(t)e^{itu}dt, \quad u \in \mathbb{R}. \tag{4.8}$$

Theorem 4.1 *Let $r > 0$, $\{a_j\}$ be an absolutely summable sequence of complex numbers, $\{\ell_j\}$ be a sequence of nonnegative, nondecreasing numbers with $\ell_j \to \infty$ as $j \to \infty$, and*

$$\tilde{K}(t) = \sum_{j=0}^{\infty} \exp(-\ell_j^2 t)a_j, \qquad \tilde{W}(t) = \sum_{j=0}^{\infty} \cos(\ell_j t)a_j.$$

Then

$$|\tilde{K}(t)| \le c_1 t^{-c_2} \exp(-r^2/t) \sum_{j=0}^{\infty} |a_j|, \quad t \in (0, 1], \tag{4.9}$$

if and only if $\tilde{W}(t) = 0$ for $0 \le t \le 2r$.

Proof Without loss of generality, we may assume that $\sum_{j=0}^{\infty} |a_j| = 1$, so that $|\tilde{W}(t)| \le 1$, $|\tilde{K}(t)| \le 1$, $t > 0$. In this proof, all the constants retain their value.

We recall the well known formula

$$\begin{aligned}\exp(-t^2/2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-u^2/2 - iut)du \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \exp(-u^2/2)\cos(ut)du, \quad t \in \mathbb{R}.\end{aligned}$$

Using this formula with $\sqrt{2t}\ell_j$ in place of t, we obtain for $t > 0$ and $j = 0, 1, \ldots,$

$$\begin{aligned}\exp(-\ell_j^2 t) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \exp(-u^2/2)\cos(u\sqrt{2t}\ell_j)du \\ &= \frac{1}{\sqrt{\pi t}} \int_0^{\infty} \exp(-u^2/(4t))\cos(\ell_j u)du. \end{aligned} \tag{4.10}$$

Since the series defining $\tilde{W}(t)$ and $\tilde{K}(t)$ converge absolutely, we may use Fubini's theorem to conclude that

$$\tilde{K}(t) = \frac{1}{\sqrt{\pi t}} \int_0^{\infty} \exp(-u^2/(4t))\tilde{W}(u)du, \quad t > 0. \tag{4.11}$$

Let $\tilde{W}(u) = 0$ for $u \in [0, 2r]$. In the case when $r^2 \le t$, we have

$$|\tilde{K}(t)| \le 1 \le e \exp(-r^2/t).$$

Next, let $r^2 > t$. Then we deduce using (4.11) that

$$|\tilde{K}(t)| \le \frac{1}{\sqrt{\pi t}} \int_{2r}^{\infty} \exp(-u^2/(4t))|\tilde{W}(u)|du \le \frac{1}{\sqrt{\pi t}} \int_{2r}^{\infty} \exp(-u^2/(4t))du$$
$$= \frac{1}{\sqrt{\pi}} \int_{r^2/t}^{\infty} u^{-1/2}e^{-u}du \le \frac{1}{\sqrt{\pi}}(r^2/t)^{-1/2}\exp(-r^2/t).$$

Since, $r^2 \ge t$, this proves (4.9).

The converse statement is much deeper. The main idea is to express the right hand side of the formula (4.11) so that it becomes a Fourier transform of a function g, supported on $[0, \infty)$. Both sides of the resulting equation can be continued analytically to the lower half plane of the complex plane $\mathbb{C}$. The estimate (4.9) together with the Fourier inversion formula can then be used to show that if ϕ is an infinitely differentiable function, supported on $[0, 4r^2]$, then $\int_{\mathbb{R}} g(u)\phi(u)du = 0$.

Let (4.9) hold. We will first extend $\tilde{K}$ as an analytic function. In this proof only, let

$$F(\zeta) := (4i\zeta)^{-1/2} \sum_{j=0}^{\infty} \exp(-\ell_j^2/(4i\zeta))a_j, \quad \zeta = \xi - i\tau \in \mathbb{C}, \tau > 0, \xi \in \mathbb{R},$$

where the principal branch of the square root is taken. The series converges uniformly and absolutely on compact subsets of the lower half plane, and $\tilde{K}(t) = F(1/(4it))t^{-1/2}$, $t > 0$. If $\zeta = \xi - i\tau$, $\tau \ge 1/4$, then $0 < \tau/(4|\zeta|^2) \le 1$, $|\exp(-\ell_j^2/(4i\zeta))| = \exp(-\ell_j^2\tau/(4|\zeta|^2))$, and we obtain from (4.9) used with $\tau/(4|\zeta|^2)$ in place of t and the fact that $|\zeta| \ge \tau$ that

$$|F(\zeta)| \le c_1(4|\zeta|)^{-1/2}(\tau/(4|\zeta|^2))^{-c_2}\exp(-4r^2|\zeta|^2/\tau)$$
$$\le c_3|\zeta|^{c_4}\tau^{-c_5}\exp(-4r^2\tau), \quad \tau \ge 1/4. \tag{4.12}$$

Next, in (4.11), we make the change of variables from u to $\sqrt{u}$, and use the resulting formula with $1/(4t)$ in place of t to obtain

$$(4t)^{-1/2}\tilde{K}(1/(4t)) = \int_0^{\infty} \exp(-ut)g(u)du, \quad t > 0,$$

where, in this proof only, $g(u) := (4\pi u)^{-1/2}\tilde{W}(\sqrt{u})$, if $u > 0$ and 0 otherwise. Writing $i\zeta$ in place of t, the left hand side of this equation is $F(\zeta)$, while the right hand side extends to an analytic function on the lower half plane. The above equation shows that these analytic functions coincide on the negative τ-axis, without the origin. Hence, the identity theorem of complex analysis implies that

$$F(\xi - i\tau) = \int_0^{\infty} e^{-u\tau}e^{-iu\xi}g(u)du, \quad \xi \in \mathbb{R}, \tau > 0. \tag{4.13}$$

Next, let $b > 0$, and ϕ be an infinitely differentiable function supported on $[0, b]$. Then (4.7) defines $\hat{\phi}$ also for all complex values of t, making it an entire function.

Moreover, an integration by parts in (4.7) shows that for any integer $R > 0$,

$$\hat{\phi}(\zeta) = (i\zeta)^{-R} \int_{\mathbb{R}} e^{-i\zeta t} \phi^{(R)}(t)dt. \tag{4.14}$$

In particular, if $\tau > 0$, $e^{-u\tau}\hat{\phi}(-\xi + i\tau)g(u)$ is absolutely integrable on $\mathbb{R} \times \mathbb{R}$ with respect to $dud\xi$, and the Fourier inversion formula holds with $e^{\tau\circ}\phi$. Since $\widehat{e^{\tau\circ}\phi}(\xi) = \hat{\phi}(\xi + i\tau)$, this formula implies that

$$e^{u\tau}\phi(u) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\phi}(\xi + i\tau)e^{iu\xi}d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\phi}(-\xi + i\tau)e^{-iu\xi}d\xi.$$

Using Fubini's theorem, we then see using (4.13) that

$$\begin{aligned}
\int_{\mathbb{R}} g(u)\phi(u)du &= \int_{\mathbb{R}} e^{-u\tau}g(u)e^{u\tau}\phi(u)du \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-u\tau}g(u)\int_{\mathbb{R}} \hat{\phi}(-\xi + i\tau)e^{-i\xi u}d\xi du \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\phi}(-\xi + i\tau)\int_{\mathbb{R}} e^{-u\tau}g(u)e^{-i\xi u}dud\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\phi}(-\xi + i\tau)F(\xi - i\tau)d\xi.
\end{aligned} \tag{4.15}$$

In this proof only, let R be the least integer greater than $2c_4$. Since $\phi^{(R)}$ is supported on $[0, b]$, (4.14) shows that

$$|\hat{\phi}(-\xi + i\tau)| \leq (\xi^2 + \tau^2)^{-R}e^{b\tau}V,$$

where, in this proof only, $V = \int_{\mathbb{R}} |\phi^{(R)}(t)|dt$. Together with (4.15) and (4.12), this yields for all $\tau \geq 1/4$,

$$\left|\int_{\mathbb{R}} g(u)\phi(u)du\right| = \left|\frac{1}{2\pi}\int_{\mathbb{R}} \hat{\phi}(-\xi + i\tau)F(\xi - i\tau)d\xi\right| \leq c_6V\tau^{-c_7}\exp(-\tau(4r^2 - b)).$$

If $b < 4r^2$, then letting $\tau \to \infty$, we see that $\int_{\mathbb{R}} g(u)\phi(u)du = 0$ for every infinitely differentiable function ϕ supported on $[0, b]$. Thus, g is supported on $[4r^2, \infty)$. In turn, this implies that $\tilde{W}$ is supported on $[2r, \infty)$. □

Corollary 4.1 *We continue the set up in Theorem 4.1, and assume that (4.9) holds. Let $G : \mathbb{R} \to \mathbb{R}$ be an even, bounded, integrable function such that $\hat{G}$ is also integrable and supported on $(-2r, 2r)$. Then*

$$\sum_{j=0}^{\infty} G(\ell_j)a_j = 0.$$

Proof Our assumptions on G imply that the Fourier inversion formula holds for G. Since G is even and real valued, so is $\hat{G}$, and we deduce from (4.8) (applied with G) that

$$G(u) = \frac{1}{\pi}\int_0^\infty \hat{G}(t)\cos(tu)dt, \quad u \in \mathbb{R}.$$

In turn, this implies that

$$\sum_{j=0}^{\infty} G(\ell_j)a_j = \frac{1}{\pi}\int_0^\infty \hat{G}(t)\tilde{W}(t)dt.$$

Since the support of $\hat{G}$ is a subset of $(-2r, 2r)$ and that of $\tilde{W}$ is a subset of $[2r, \infty)$, the last integral is equal to 0. □

4.2 Summability Kernels

In the remainder of this section, we will assume that $H : \mathbb{R} \to \mathbb{R}$ is a fixed function as in Theorem 2.1. Having proved the equivalence of our conditions on the heat kernels with bounds on the sum of squares of $|\phi_k|$ and $|\psi_k|$ as well as the finite speed of wave propagation, the proofs in this section are nearly the same as that of [24, Theorem 4.1]. We sketch them to ensure that they continue to be valid also with the generalization made to nonorthogonal pair of systems.

Let $V : \mathbb{R} \to \mathbb{R}$ be chosen such that V is an even function, $\hat{V}$ is an infinitely differentiable function, $\hat{V}(\omega) = 1$ if $|\omega| \le 1/2$, and $\hat{V}(\omega) = 0$ if $|\omega| \ge 1$. Then for every $Y > 0$, there exists $H_Y : \mathbb{R} \to \mathbb{R}$ such that

$$\hat{H}_Y(\omega) = \hat{H}(\omega)\hat{V}(\omega/Y), \quad \omega \in \mathbb{R}. \tag{4.16}$$

In the "time domain", one has

$$H_Y(t) = Y\int_{\mathbb{R}} H(u)V(Y(t-u))du. \tag{4.17}$$

We recall from [30, Sect. 5.2.2, eqn. (4), Sect. 4.2, eqn. (15)] that

$$\max_{t\in\mathbb{R}} |H^{(m)}(t) - H_Y{}^{(m)}(t)| \le \frac{c}{Y^{S-m}} \max_{t\in\mathbb{R}} |H^{(S)}(t)|, \quad m = 0, 1, \ldots, S. \tag{4.18}$$

We will write $K = (A_1 + A_2)/2$, where A_1, A_2 are defined in (2.5) (cf. (2.7)).

The following lemma is a simple variation of [24, Lemma 6.1].

Lemma 4.1

(a) *If $G : \mathbb{R} \to \mathbb{R}$ is a bounded function, $\{a_j\}$ is a bounded sequence, then we have for any $C > 0$, $L \ge 1$,*

$$\left|\sum_{\ell_j \le CL} G(\ell_j/L)a_j\phi_j(x)\overline{\psi_j(y)}\right| \le c(CL)^K \sup_{t\in[0,C]} |G(t)| \max_{\ell_j\le CL} |a_j|, \quad x, y \in \mathbb{X}. \tag{4.19}$$

(b) *Let G be a continuous, integrable, even, real valued function on $\mathbb{R}$, vanishing at infinity, such that the Fourier transform $\hat{G}$ is also integrable. Let f_1 (respectively, f_2) be a function for which (2.3) holds with $\{\phi_k\}$ (respectively, $\{\psi_k\}$). Then for every $L > 0$,*

$$\sum_j G(\ell_j/L)\langle f_1, \phi_j\rangle\overline{\langle f_2, \psi_j\rangle} = \frac{L}{\pi}\int_0^\infty \hat{G}(Lt)W(t, f_1, f_2)dt. \tag{4.20}$$

(c) *For any $x \in \mathbb{X}$, $r > 0$, and a nonincreasing function $g : [0, \infty) \to [0, \infty)$,*

$$L^\alpha \int_{\Delta(x,r)} g(L\rho(x, y))d\mu(y) \le c\int_{rL/2}^\infty g(v)v^{\alpha-1}dv. \tag{4.21}$$

Proof Part (a) is a simple consequence of (2.7) and the Schwarz inequality. The assumptions on G imply that the Fourier inversion formula holds; i.e.,

$$G(u/L) = \frac{1}{\pi}\int_0^\infty \hat{G}(v)\cos(vu/L)dv = \frac{L}{\pi}\int_0^\infty \hat{G}(Lt)\cos(tu)dt.$$

Since (2.3) holds for f_1, f_2, equation (4.20) follows by an application of Fubini's theorem. Part (c) is proved in [24] using (2.2), where the symbol K is used in place of α. □

Lemma 4.2 *Let $\{a_j\}$ be a bounded sequence of complex numbers. For $x, y \in \mathbb{X}$, $Y, L \ge 1/2$, $J \ge L$, and integer $N > K - 1$, we have*

$$\left|\sum_{\ell_j \ge J} H_Y(\ell_j/L)a_j\phi_j(x)\overline{\psi_j(y)}\right| \le c(N)J^K Y^{-N}(L/J)^{N+1}\max_j |a_j|. \tag{4.22}$$

In particular, the series $\sum_{j=0}^\infty H_Y(\ell_j/L)a_j\phi_j(x)\overline{\psi_j(y)}$ converges uniformly on $\mathbb{X} \times \mathbb{X}$ to a continuous and bounded function, and

$$\left|\sum_{\ell_j \ge 2L} H_Y(\ell_j/L)a_j\phi_j(x)\overline{\psi_j(y)}\right| \le c(N)L^K Y^{-N}\max_j |a_j|. \tag{4.23}$$

Proof The proof is the same as that of [24, Lemma 6.2], except that the function s in that proof should be replaced by $s(u) = \sum_{\ell_j \le u} a_j\phi_j(x)\overline{\psi_j(y)}$, and we use (2.7) and the Schwarz inequality to conclude that $|s(u)| \le cu^K$. □

The following lemma is proved using Lemma 4.2 and Lemma 4.1(a) as in the proof of [24, Lemma 6.3]. The proof being verbatim the same, except for substituting $\phi_j(x)\overline{\psi_j(y)}$ for $\phi_j(x)\phi_j(y)$, we omit it.

Lemma 4.3 *Let* $x, y \in \mathbb{X}$, $Y \geq 1/2$. *Let* $\{a_j\}$ *be a bounded sequence of complex numbers with* $\max_j |a_j| \leq 1$. *Then*

$$\left|\sum_{j=0}^{\infty}(H(\ell_j/L) - H_Y(\ell_j/L))a_j\phi_j(x)\overline{\psi_j(y)}\right| \leq cL^K Y^{-S}. \tag{4.24}$$

We are now in a position to prove Theorem 2.1.

Proof of Theorem 2.1 In light of (4.19), we may assume that $r := (\rho(x, y)/2\kappa^2) \geq 1/L$. Let $Y = \sqrt{\kappa_5}Lr/\kappa^2$, where κ_5 is the constant appearing in (2.6). Next, let $f, g \in L^1$ be supported on $B(x, \rho(x, y)/8\kappa(1 + \kappa))$, $B(y, \rho(x, y)/8\kappa(1 + \kappa))$ respectively, $\|f\|_1 = \|g\|_1 = 1$, and $0 < \epsilon < 1$ be arbitrary. Then there exist continuous functions f_1, g_1 such that $\|f_1 - f\|_1 \leq \epsilon$, $\|g_1 - g\|_1 \leq \epsilon$. Multiplying f_1 by a continuous function having range in [0, 1], equal to 1 on $B(x, \rho(x, y)/8\kappa(1 + \kappa))$ and 0 outside $B(x, \rho(x, y)/4\kappa(1 + \kappa))$, we obtain a continuous function $f_2 \in L^1 \cap L^\infty$, such that $\|f - f_2\|_1 \leq 2\epsilon$. Similarly, we obtain a continuous function $g_2 \in L^1 \cap L^\infty$, such that $\|g - g_2\|_1 \leq 2\epsilon$, and g_2 is supported on $B(y, \rho(x, y)/4\kappa(1 + \kappa))$. We may then obtain $f_3 \in \mathcal{D}(\{\phi_k\})$ supported on $B(x, \rho(x, y)/2\kappa(1 + \kappa))$ (respectively, $g_3 \in \mathcal{D}(\{\psi_k\})$, supported on $B(y, \rho(x, y)/2\kappa(1 + \kappa)))$ such that $\|f - f_3\|_1 \leq 3\epsilon$, $\|g - g_3\|_1 \leq 3\epsilon$. Now, (4.19) implies that

$$\begin{aligned}
&\left|\sum_j H(\ell_j/L)\langle f, \phi_j\rangle\overline{\langle g, \psi_j\rangle} - \sum_j H(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \\
&\quad = \left|\sum_j H(\ell_j/L)\langle f, \phi_j\rangle\overline{\langle g, \psi_j\rangle} - \sum_j H(\ell_j/L)\langle f, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \\
&\qquad + \left|\sum_j H(\ell_j/L)\langle f, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle} - \sum_j H(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \\
&\quad \leq cL^K\|g - g_3\|_1 + cL^K\|f - f_3\|_1 \leq c\epsilon L^K.
\end{aligned} \tag{4.25}$$

Next, using Lemma 4.3, we conclude that

$$\begin{aligned}
&\left|\sum_j H(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle} - \sum_j H_Y(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \\
&\quad \leq cL^K Y^{-S}\|f_3\|_1\|g_3\|_1 \leq cL^K Y^{-S}.
\end{aligned} \tag{4.26}$$

Now, the sequences $\{\langle f_3, \phi_j\rangle\}$ and $\{\langle g_3, \psi_j\rangle\}$ are square summable and the distance between the supports of f_3 and g_3 exceeds $\rho(x, y)/(2\kappa^2)$. So, (2.6) implies that

$$\left|\sum_{j=0}^{\infty}\exp(-\ell_j^2 t)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \leq \kappa_4 t^{-A_3}\exp(-\kappa_5\rho(x, y)^2/(4\kappa^4 t)), \quad t \in (0, 1];$$

i.e., (4.9) holds with the absolutely summable sequence $\{\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\}$ as $\{a_j\}$, and $\sqrt{\kappa_5}r$ in place of r. Since the support of $\widehat{H_Y(\circ/L)}$ is contained in $[-Y/L, Y/L] = [-2\sqrt{\kappa_5}r, 2\sqrt{\kappa_5}r]$, Corollary 4.1 implies that

$$\sum_j H_Y(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle} = 0.$$

Thus, we see from (4.26) that

$$\left|\sum_j H(\ell_j/L)\langle f_3, \phi_j\rangle\overline{\langle g_3, \psi_j\rangle}\right| \le cL^K Y^{-S}.$$

Together with (4.25) and the fact that ϵ is arbitrary, this implies that

$$\left|\sum_j H(\ell_j/L)\langle f, \phi_j\rangle\overline{\langle g, \psi_j\rangle}\right| \le cL^K Y^{-S}.$$

Since f and g are arbitrary functions in L^1 supported near x, y respectively, and ϕ_j, ψ_j are continuous, this implies (2.10).

The proof of the remaining parts of this theorem using (2.10) and (2.2) is standard, see for example the proof of [24, Theorem 4.2, Proposition 2.1]. Thus, we observe using (4.21) with $g(v) = (\max(1, v))^{-S}$ that for $r > 0$,

$$L^\alpha \int_{\Delta(x,r)} (\max(1, L\rho(x, y)))^{-S} d\mu(y) \le c\int_{rL/2}^\infty g(v)v^{\alpha-1}dv \le c(\max(1, Lr))^{\alpha-S}. \tag{4.27}$$

Therefore, (2.10) implies that

$$\int_{\Delta(x,r)} \left|\sum_k H(\ell_k/(2L))\phi_k(x)\overline{\psi_k(y)}\right| d\mu(y) \le cL^{K-\alpha}(\max(1, Lr))^{\alpha-S}. \tag{4.28}$$

The estimates (4.19) and (2.2) imply that

$$\int_{B(x,r)} \left|\sum_k H(\ell_k/(2L))\phi_k(x)\overline{\psi_k(y)}\right| d\mu(y) \le cL^K r^\alpha.$$

Choosing $r \sim 1/L$ in the above two estimates, we deduce (2.11). The estimate (2.12) follows from (2.11) by an application of the Riesz–Thorin interpolation theorem. □

Proof of Theorem 2.2 In this proof only, we choose $H : \mathbb{R} \to [0, 1]$ to be a fixed, even, infinitely often differentiable function, nonincreasing on $[0, \infty)$, such that $H(t) = 1$ if $|t| \le 1/2$, and $H(t) = 0$ if $|t| > 1$. If $P \in \Pi_L$, then it is easy to verify using the orthonormality of ϕ_k that

$$P(y) = \int_{\mathbb{X}} P(x)\sum_k H(\ell_k/(2L))\overline{\phi_k(x)}\phi_k(y)d\mu(x), \tag{4.29}$$

and

$$(TP)(y) = \int_{\mathbb{X}} P(x) \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) d\mu(x)$$
$$= \sum_{k=0}^{\infty} H(\ell_k/(2L)) \langle P, \phi_k \rangle \psi_k(y). \tag{4.30}$$

The estimate (2.14) follows from (4.30) and (2.12). Next, let $C_{3/2}$ be an open set, containing the closure of C_1, and with the closure of $C_{3/2}$ being a subset of C_2. We can also arrange that $\inf_{x\in\mathbb{X}\setminus C_{3/2}} \rho(x, C_1) \ge d/2$; i.e., $\mathbb{X} \setminus C_{3/2} \subseteq \Delta(y, d/2)$ for each $y \in C_1$, and $C_1 \subseteq \Delta(x, \delta/2)$ for each $x \in \mathbb{X} \setminus C_{3/2}$. Using (4.28) (with different variable names), we deduce that

$$\sup_{y\in C_1} \int_{\mathbb{X}\setminus C_{3/2}} \left| \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) \right| d\mu(x)$$
$$\le \sup_{y\in C_1} \int_{\Delta(y,d/2)} \left| \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) \right| d\mu(x)$$
$$\le c \frac{L^{K-\alpha}}{(\max(1, Ld))^S}. \tag{4.31}$$

Let $\phi : \mathbb{X} \to [0, 1]$ be a continuous function, equal to 1 on $C_{3/2}$, and 0 outside of C_2. We consider an operator defined on L^1 (in this proof only) by

$$Uf(y) := \int_{\mathbb{X}} f(x)(1 - \phi(x)) \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) d\mu(x)$$
$$= \int_{\mathbb{X}\setminus C_{3/2}} f(x)(1 - \phi(x)) \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) d\mu(x).$$

In view of (4.31), and the fact that $C_1 \subseteq \Delta(x, \delta/2)$ for each $x \in \mathbb{X} \setminus C_{3/2}$, we conclude that for $p = 1, \infty$,

$$\|Uf\|_{p,C_1} \le c \frac{L^{K-\alpha}}{(\max(1, Ld))^S} \|f\|_p. \tag{4.32}$$

An application of the Riesz–Thorin theorem shows that (4.32) holds also for all p, $1 \le p \le \infty$.

In view of (4.30)

$$TP(y) = \int_{\mathbb{X}} P(x)\phi(x) \sum_k H(\ell_k/(2L)) \overline{\phi_k(x)} \psi_k(y) d\mu(x) + UP(y)$$
$$= \sum_k H(\ell_k/(2L)) \langle P\phi, \phi_k \rangle \psi_k(y) + UP(y).$$

Therefore, using (2.12) and (4.32), we conclude that

$$\begin{aligned}\|TP\|_{p,C_1} &\le \left\| \sum_k H(\ell_k/(2L))\langle P\phi, \phi_k\rangle \psi_k \right\|_p + \|UP\|_{p,C_1} \\ &\le cL^{K-\alpha}\|P\phi\|_p + c\frac{L^{K-\alpha}}{(\max(1, Ld))^S}\|P\|_p \\ &\le cL^{K-\alpha}\{\|P\|_{p,C_2} + (\max(1, Ld))^{-S}\|P\|_p\}.\end{aligned}$$

This proves (2.13). □

4.3 Quadrature Formulas

First, we will prove Theorem 3.2. Together with the Hahn–Banach theorem (respectively, Krein–Rutman theorem), this will lead to part (a) (respectively, part (b)) of Theorem 3.1. The definition of the mesh norm implies that $\mathbb{X} = \bigcup_{x\in\mathcal{C}} B(x, \delta_{\mathcal{C}})$. In order to prove (3.5), we need to also consider the balls $B(x, 2\delta_{\mathcal{C}})$, and the disjoint balls $B(x, \delta_{\mathcal{C}}/(2\kappa_6))$ (cf. (3.3)). In the following lemmas, we obtain the necessary interconnections among these and the integrals of diffusion polynomials on these.

Lemma 4.4 *Let $\mathcal{C}$ be a finite set for which* (3.3) *holds, and $\delta_{\mathcal{C}} \le 2\kappa_6$. If $c_1 > 0$, the number of balls $B(x, c_1\delta_{\mathcal{C}})$, $x \in \mathcal{C}$, which intersect each other is bounded from above by a constant, dependent only on c_1, but independent of $\delta_{\mathcal{C}}$.*

Proof In this proof, let $\delta := \delta_{\mathcal{C}}$. Let $y_1, \ldots, y_m \in \mathcal{C}$ and $y \in \bigcap_{k=1}^m B(y_k, c_1\delta)$. Then $B(y, \delta) \subset \bigcap_{k=1}^m B(y_k, (1 + c_1)\delta)$. Since $q_{\mathcal{C}} \ge \delta/\kappa_6$, the balls $B(y_k, \delta/(2\kappa_6))$ are pairwise disjoint, and their union is a subset of $B(y, (1 + c_1)\delta)$. Therefore, (3.9) implies that

$$\begin{aligned}\mu(B(y,\delta)) &\le \min_{1\le k\le m} \mu(B(y_k, (1 + c_1)\delta)) \le \frac{1}{m}\sum_{k=1}^m \mu(B(y_k, (1 + c_1)\delta)) \\ &\le \frac{c_2}{m}\sum_{k=1}^m \mu(B(y_k, \delta/(2\kappa_6))) = \frac{c_2}{m}\mu\left(\bigcup_{k=1}^m B(y_k, \delta/(2\kappa_6))\right) \\ &\le \frac{c_2}{m}\mu(B(y, (1 + c_1)\delta)) \le \frac{c_3}{m}\mu(B(y, \delta)).\end{aligned}$$

Thus, $m \le c_3$. □

In the sequel, let $\mathcal{C} = \{x_1, \ldots, x_M\}$, $X_k = B(x_k, \delta_{\mathcal{C}})$, $\tilde{X}_k = B(x_k, 2\delta_{\mathcal{C}})$, $k = 1, \ldots, M$.

Lemma 4.5 *Let T be an operator as in Theorem* 2.2. *For $L > 0$, $P \in \Pi_L$,*

$$\sum_{k=1}^M \mu(X_k)\|TP\|_{\infty,\tilde{X}_k} \le cL^{K-\alpha}\{(\delta_{\mathcal{C}}L)^\alpha + \min(1, (\delta_{\mathcal{C}}L)^{\alpha-S})\}\|P\|_1. \quad (4.33)$$

We note that the constant c in this lemma depends on T only through the constants $\kappa, \kappa_1, \ldots, \kappa_6, A_1, A_2$ in (2.2), (2.5) and (2.6) with $\psi_k = T\phi_k$.

Proof of Lemma 4.5 In this proof, we will again write $\delta = \delta_{\mathcal{C}}$, and recall that $\delta \le 1/6$. In this proof only, we let $H : \mathbb{R} \to [0, 1]$ to be a fixed, even, infinitely often differentiable function, nonincreasing on $[0, \infty)$, such that $H(t) = 1$ if $|t| \le 1/2$, and $H(t) = 0$ if $|t| > 1$, $T\phi_k = \psi_k$, and

$$\Psi_{2L}(x, y) := \sum_k H\left(\frac{\ell_k}{2L}\right)\phi_k(x)\overline{\psi_k(y)}, \quad x, y \in \mathbb{X}, L > 0.$$

In view of Theorem 2.1, we have

$$|\Psi_{2L}(x, z)| \le \frac{cL^K}{\max(1, L\rho(x, z))^S}, \quad x, z \in \mathbb{X}, L > 0. \tag{4.34}$$

Let $x \in \mathbb{X}$, $\mathcal{I} := \{j : \rho(x, \tilde{X}_j) \ge 5\delta\}$, $\mathcal{I}' := \{1, \ldots, M\} \setminus \mathcal{I}$. If $j \in \mathcal{I}$, $z_j \in \tilde{X}_j$ is chosen so that $\rho(x, z_j) = \rho(x, \tilde{X}_j)$, then for any $y \in \tilde{X}_j$, we have

$$|\rho(x, y) - \rho(x, \tilde{X}_j)| = |\rho(x, y) - \rho(x, z_j)| \le \rho(y, z_j) \le 4\delta \le (4/5)\rho(x, \tilde{X}_j).$$

Therefore, for any $y \in \tilde{X}_j$,

$$\delta \le (1/5)\rho(x, \tilde{X}_j) \le \rho(x, y) \le (9/5)\rho(x, \tilde{X}_j). \tag{4.35}$$

Consequently, we conclude from (4.34) that for any $j \in \mathcal{I}$, $y \in X_j$, $z \in \tilde{X}_j$,

$$|\Psi_{2L}(x, z)| \le \frac{cL^K}{\max(1, L\rho(x, \tilde{X}_j))^S} \le \frac{cL^K}{\max(1, L\rho(x, y))^S}. \tag{4.36}$$

We will write again $g(t) = 1/(\max(1, t))^S$, $t \ge 0$. In view of Lemma 4.4, at most finitely many of the balls $\tilde{X}_j$ can intersect each other. Hence, using (4.27), we deduce that

$$\begin{aligned}\sum_{j\in\mathcal{I}} \mu(X_j) \max_{z\in\tilde{X}_j} |\Psi_{2L}(x, z)| &\le cL^K \sum_{j\in\mathcal{I}} \int_{X_j} g(L\rho(x, y))d\mu(y) \\ &\le cL^K \int_{\Delta(x,\delta)} g(L\rho(x, y))d\mu(y) \\ &\le cL^{K-\alpha} \max(1, \delta L)^{\alpha-S}. \end{aligned} \tag{4.37}$$

Lemma 4.4 shows further that $|\mathcal{I}'| \le c$. Using $|\Psi_{2L}(x, z)| \le cL^K$ for all $x, z \in \mathbb{X}$, we obtain

$$\sum_{j\in\mathcal{I}'} \mu(X_j) \max_{z\in\tilde{X}_j} |\Psi_{2L}(x, z)| \le cL^K \sum_{j\in\mathcal{I}'} \mu(X_j) \le cL^K \mu\left(\bigcup_{j\in\mathcal{I}'} X_j\right) \le cL^{K-\alpha}(\delta L)^\alpha. \tag{4.38}$$

From (4.37), (4.38), we conclude that

$$\sum_{j=1}^{M}\mu(X_j)\max_{z\in\tilde{X}_j}|\Psi_{2L}(x,z)|\le cL^{K-\alpha}\{(\delta L)^\alpha+\min(1,(\delta L)^{\alpha-S})\},\quad x\in\mathbb{X}. \tag{4.39}$$

Next, let $P\in\Pi_L$. Since

$$(TP)(z)=\int_{\mathbb{X}}P(x)\Psi_{2L}(z,x)d\mu(x),\quad z\in\tilde{X}_j,$$

(4.39) implies that

$$\begin{aligned}\sum_{k=1}^{M}\mu(X_k)\|TP\|_{\infty,\tilde{X}_k}&\le\int_{\mathbb{X}}|P(x)|\left\{\sum_{k=1}^{M}\mu(X_k)\max_{z\in\tilde{X}_k}|\Psi_{2L}(x,z)|\right\}d\mu(x)\\&\le cL^{K-\alpha}\{(\delta L)^\alpha+\min(1,(\delta L)^{\alpha-S})\}\|P\|_1,\end{aligned}$$

which is (4.33). □

Let F be any vector field on $\mathbb{X}$ with $|||F|||_x\equiv 1$ (see Appendix for definition). We choose T in the above lemma defined by $Tf(x)=\langle\nabla f(x),F(x)\rangle_x$. In view of the remarks after Theorem 2.2, the constants $\kappa_1,\ldots,\kappa_5$, A_1,A_2 are independent of F. Moreover, we may choose $K=\alpha+1$. Consequently, for *any* vector field F with $|||F|||_x\equiv 1$, we have

$$\begin{aligned}&\sum_{k=1}^{M}\mu(X_k)\|\langle\nabla P(x),F(x)\rangle_x\|_{\infty,\tilde{X}_k}\\&\quad\le cL\{(\delta L)^\alpha+\min(1,(\delta L)^{\alpha-S})\}\|P\|_1,\quad P\in\Pi_L.\end{aligned}$$

If $P\in\Pi_L$, this implies that

$$\sum_{k=1}^{M}\mu(X_k)\||||\nabla P|||_\circ\|_{\infty,\tilde{X}_k}\le cL\{(\delta L)^\alpha+\min(1,(\delta L)^{\alpha-S})\}\|P\|_1,\quad P\in\Pi_L. \tag{4.40}$$

Proof of Theorem 3.2 In this proof, we continue to write δ in place of $\delta_{\mathcal{C}}$. We also find it convenient to let the various constants retain their value, within this proof only. Let $P\in\Pi_L$, $k=1,\ldots,M$, $z\in X_k$. Since $\mathbb{X}$ is compact and connected, there is a geodesic γ from x_k to z with length equal to $\rho(z,x_k)$ [3, Corollary 7.11]. Necessarily, $\gamma\subset X_k$. This geodesic can be covered by overlapping normal neighborhoods of finitely many points. Then $P(z)-P(x_k)$ can be expressed as a sum of line integrals of inner products of ∇P with the tangent fields for the parts of γ in each neighborhood. Since only finitely many of the X_k's intersect each other, $\sum_{k=1}^{M}\mu(X_k)\le c\mu(\mathbb{X})=c$.

Therefore, using (4.40), we deduce that

$$\sum_{k=1}^{M}\mu(X_k)\max_{z\in X_k}|P(z)-P(x_k)| \le \delta\sum_{k=1}^{M}\mu(X_k)\|\|\!|\nabla P|\!\|_\circ\|_{\infty,\tilde{X}_k} \le c_1\delta L\{(\delta L)^\alpha+\min(1,(\delta L)^{\alpha-S})\}\|P\|_1. \quad (4.41)$$

This leads to (3.5) by choosing $\delta L \le \eta\min(1, 1/(2c_1))$.

Next, let $Y_1 = X_1$, $Y_k = X_k \setminus (\bigcup_{j=1}^{k-1} X_j)$, $k = 2,\dots,M$. Then the sets Y_k's are pairwise disjoint, $\mathbb{X} = \bigcup_{k=1}^{M} Y_k$, $B(x_k, q_{\mathcal{C}}/2) \subset Y_k \subset X_k$, $k = 1,\dots,M$. Therefore, (3.9) and (3.3) imply that $\mu(Y_k) \sim \mu(X_k)$, $k = 1,\dots,M$. Thus, (4.41) may be rewritten in the form

$$\sum_{k=1}^{M}\int_{Y_k}|P(z)-P(x_k)|d\mu(z) \le \sum_{k=1}^{M}\mu(Y_k)\max_{z\in X_k}|P(z)-P(x_k)| \le c_2\delta L(1+(\delta L)^\alpha)\|P\|_1. \quad (4.42)$$

Choosing $\delta L \le \eta\min(1, 1/(2c_2))$ yields

$$\left|\sum_{k=1}^{M}\mu(Y_k)|P(x_k)| - \int_{\mathbb{X}}|P(z)|d\mu(z)\right| = \left|\sum_{k=1}^{M}\int_{Y_k}|P(x_k)|d\mu(z) - \sum_{k=1}^{M}\int_{Y_k}|P(z)|d\mu(z)\right| \le \sum_{k=1}^{M}\int_{Y_k}|P(z)-P(x_k)|d\mu(z) \le \eta\|P\|_1. \quad (4.43)$$

Since $\mu(Y_k) \sim \mu(X_k)$, $k = 1,\dots,M$, this leads to (3.6).

Next, we prove (3.7). Let $P(x_k) \ge 0$ for $k = 1,\dots,M$. We choose $\eta = 1/3$, $\delta L \le \eta\min(1, 1/(2c_2))$. Then (4.43) implies that

$$(2/3)\|P\|_1 \le \sum_{k=1}^{M}\mu(Y_k)|P(x_k)| = \sum_{k=1}^{M}\mu(Y_k)P(x_k) \le (4/3)\|P\|_1. \quad (4.44)$$

Using (4.43) again, we obtain

$$\int_{\mathbb{X}}P(z)d\mu(z) = \sum_{k=1}^{M}\int_{Y_k}P(z)d\mu(z) \ge \sum_{k=1}^{M}\mu(Y_k)P(x_k) - (1/3)\|P\|_1 \ge (1/2)\sum_{k=1}^{M}\mu(Y_k)P(x_k).$$

Since $\mu(Y_k) \sim \mu(X_k)$, $k = 1,\dots,M$, this leads to (3.7). □

We are now in a position to prove part (a) of Theorem 3.1, using the technique developed in [28]. We sketch a proof of the sake of completeness.

Proof of Theorem 3.1(a) We adopt some notations, to be valid in this proof only. For $\mathbf{y} \in \mathbb{R}^M$, let $\|\mathbf{y}\| = \sum_{k=1}^{M} \mu(B(x_k, \delta_{\mathcal{C}}))|y_k|$; the dual norm being $\|\mathbf{y}\|^* = \max_{1\le k\le M} |y_k|/\mu(B(x_k, \delta_{\mathcal{C}}))$. We consider an operator $\mathcal{S} : \Pi_L \to \mathbb{R}^M$ given by $\mathcal{S}(P) = (P(x_1), \ldots, P(x_M))$, and let W denote the range of $\mathcal{S}$. In view of (3.6), the operator $\mathcal{S}$ is injective, and hence, invertible as an operator from Π_L to W. We now define a functional x^* on W by

$$x^*(\mathbf{y}) = \int_{\mathbb{X}} \mathcal{S}^{-1}(\mathbf{y})d\mu, \quad \mathbf{y} \in W.$$

The norm of this functional can be estimated using (3.6) as follows:

$$\sup_{\mathbf{y}\in W} |x^*(\mathbf{y})|/\|\mathbf{y}\| = \sup_{P\in\Pi_L} \frac{\int_{\mathbb{X}} P d\mu}{\|(P(x_1), \ldots, P(x_M))\|} \le c.$$

In view of the Hahn–Banach theorem, the functional x^* admits an extension to $\mathbb{R}^M$ with the same norm. We may identify this extended functional with a vector $(w_1, \ldots, w_M)$; i.e., the extended functional is given by $\mathbf{y} \mapsto \sum_{k=1}^{m} w_k y_k$, $\mathbf{y} \in \mathbb{R}^M$. The fact that this extends the functional x^* means that for $P \in \Pi_L$,

$$\sum_{k=1}^{M} w_k P(x_k) = x^*((P(x_1), \ldots, P(x_M)) = \int_{\mathbb{X}} P d\mu,$$

which is (3.4). The norm of this functional is on one hand, the dual norm $\|(w_1, \ldots, w_M)\|^*$, and on the other hand, bounded by c. This shows that $|w_k| \le c\mu(B(x_k, \delta_{\mathcal{C}}))$. □

In order to prove Theorem 3.1(b), we recall an abstract quadrature formula in the setting of general finite dimensional spaces (cf. [25, Theorem 3.2.1]).

Proposition 4.2 *Let* $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ *be a finite dimensional normed linear space,* $(\mathcal{V}^*, \|\cdot\|_{\mathcal{V}^*})$ *be its dual space,* $\mathcal{Z} = \{x_1^*, \ldots, x_M^*\} \subset \mathcal{V}^* \setminus \{0\}$, *and* $x^* \in \mathcal{V}^*$. *Suppose the operator* $x \mapsto (x_1^*(x), \ldots, x_M^*(x))$, $x \in \mathcal{V}$, *is one-to-one, and the following two conditions are satisfied*: (1) *If* $x \in \mathcal{V}$ *and* $x_\ell^*(x) \ge 0$ *for* $\ell = 1, \ldots, M$, *then* $x^*(x) \ge 0$, *and* (2) *there exists some* $x_0 \in \mathcal{V}$ *such that* $x_\ell^*(x_0) > 0$ *for* $\ell = 1, \ldots, M$. *Then there exist nonnegative numbers* W_ℓ, $\ell = 1, \ldots, M$ *such that*

$$x^*(x) = \sum_{\ell=1}^{M} W_\ell x_\ell^*(x), \quad x \in \mathcal{V}. \tag{4.45}$$

Proof of Theorem 3.1(b) We use Proposition 4.2 with Π_L in place $\mathcal{V}$, and the point evaluation functionals $x_k^*(P) = P(x_k)$, $k = 1, \ldots, M$. For x^*, we take the functional

$$x^*(P) = \int_{\mathbb{X}} P(z)d\mu(z) - (c_3/2)\sum_{k=1}^{M} \mu(X_k)P(x_k),$$

where c_3 is the constant appearing in (3.7). The estimate (3.7) then shows that the condition (1) in Proposition 4.2 is satisfied. The condition (2) is satisfied by the element $P_0 \equiv 1$. Then Proposition 4.2 implies the existence of $W_k \geq 0$, $k = 1, \ldots, M$, such that (4.45) is satisfied. We define $w_{x_k}^+ := w_k^+ := W_k + (c_3/2)\mu(X_k)$. Then it is easy to verify that (3.4) is satisfied as claimed in Theorem 3.1(b). Moreover, it is clear that $w_k^+ \geq (c_3/2)\mu(X_k)$, $k = 1, \ldots, M$. □

Acknowledgements We would like to thank Professor J. Korevaar for sending us [22], and his encouragement. We thank also the referees for their careful reading and suggestions for the improvement in our original draft.

Appendix

For the convenience of the reader, we include in this appendix the definitions of some of the terminology regarding Riemannian manifolds in this paper. We will avoid very technical details, which can be found in such standard texts as [3, 8, 9, 31]. The material in this appendix (and the notation) is based mostly on [9].

In this appendix, let $n \geq 1$ be a fixed integer. A *differentiable manifold* of dimension n is a set $\mathbb{X}$ and a family of injective mappings $\mathbf{x}_\alpha : U_\alpha \subset \mathbb{R}^n \to \mathbb{X}$ of open sets U_α into $\mathbb{X}$ such that (i) $\cup_\alpha \mathbf{x}_\alpha(U_\alpha) = \mathbb{X}$, (ii) for any pair α, β, with $\mathbf{x}_\alpha(U_\alpha) \cap \mathbf{x}_\beta(U_\beta) = W$ being nonempty, the sets $\mathbf{x}_\alpha^{-1}(W)$ and $\mathbf{x}_\beta^{-1}(W)$ are open subsets of $\mathbb{R}^n$, and the mapping $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha$ is (infinitely) differentiable on $\mathbf{x}_\alpha^{-1}(W)$. (iii) The family (*atlas*) $\mathcal{A}_\mathbb{X} = \{(U_\alpha, \mathbf{x}_\alpha)\}$ is maximal relative to the conditions (i) and (ii). The pair $(U_\alpha, \mathbf{x}_\alpha)$ (respectively, $\mathbf{x}_\alpha$) with $p \in \mathbf{x}_\alpha(U_\alpha)$ is called a parametrization or coordinate chart (respectively, a system of coordinates) of $\mathbb{X}$ around p, and $\mathbf{x}_\alpha(U_\alpha)$ is called a coordinate neighborhood of p. In the sequel, the term differentiable will mean infinitely many times differentiable. We assume also that $\mathbb{X}$ is Hausdorff and has a countable basis as a topological space.

Intuitively, one thinks of a differentiable manifold as a surface in an ambient Euclidean space. The abstract definition above is intended to overcome the technical need for the ambient space. For all applications of our theory that we can imagine, and in particular, for an intuitive comprehension of our paper, there is no loss in thinking of a manifold as a surface. Moreover, a theorem of Whitney [9, p. 30] provides a further justification of such a viewpoint.

Let $\mathbb{X}, \mathbb{Y}$ be two differentiable manifolds with dimension n and m respectively. A mapping $f : \mathbb{X} \to \mathbb{Y}$ is called differentiable on an open set $W \subseteq \mathbb{X}$ if there exist for every $p \in W$ coordinate charts $(U, \mathbf{x}) \in \mathcal{A}_\mathbb{X}$, $(V, \mathbf{y}) \in \mathcal{A}_\mathbb{Y}$ with $p \in U$, $f(U) \subseteq V$ such that $\mathbf{y}^{-1} \circ f \circ \mathbf{x}$ is a C^∞ function. In particular, a *curve* in $\mathbb{X}$ is a differentiable mapping from an interval in $\mathbb{R}$ to $\mathbb{X}$. The restriction of a curve γ to a compact subinterval $[a, b]$ of I is called a curve segment, joining $\gamma(a)$ to $\gamma(b)$. We may define a piecewise differentiable curve on a manifold $\mathbb{X}$ in an obvious manner.

If $p \in \mathbb{X}$, $\epsilon > 0$, and $\gamma : (-\epsilon, \epsilon) \to \mathbb{X}$ is a curve with $p = \gamma(0)$, then the *tangent vector* to γ at $\gamma(t_0)$ is defined to be the functional $\gamma'(t)$ acting on the class of all differentiable $f : \mathbb{X} \to \mathbb{R}$ by

$$\gamma'(t) f = \left. \frac{d(f \circ \gamma)}{dt} \right|_{t=t_0}.$$

The set of all such functionals $\gamma'(0)$ defines a vector space, called the *tangent space* of $\mathbb{X}$ at p, denoted by $T_p\mathbb{X}$. Let $(U, \mathbf{x})$ be a coordinate chart such that $0 \in U$ and $p = \mathbf{x}(0)$, and for $j = 1, \ldots, n$, $X_j(p)$ be the tangent vector at p to the coordinate curve $x_j \to (0, \ldots, x_j, 0, \ldots, 0)$. Then $\{X_j(p)\}$ is a basis for $T_p\mathbb{X}$. In particular, the dimension of $T_p\mathbb{X}$ is n. The set $\{(p, v) : p \in \mathbb{X}, v \in T_p\mathbb{X}\}$ is called the *tangent bundle* of $\mathbb{X}$, and can be endowed with the structure of a differentiable manifold of dimension $2n$. A *vector field* F on $\mathbb{X}$ is a mapping that assigns to each $p \in \mathbb{X}$ a vector $F(p) \in T_p(\mathbb{X})$ such that for every differentiable f on $\mathbb{X}$, the mapping $p \mapsto F(p)g$ is differentiable. If G is another vector field, we may apply $G(p)$ to this mapping, obtaining thereby a second order vector field $G \circ F$. A derivative of higher order can be defined similarly.

A *Riemannian metric* on a differentiable manifold $\mathbb{X}$ is given by a scalar product $\langle \circ, \circ \rangle_p$ on each $T_p\mathbb{X}$ which depends smoothly on the base point p, i.e. the function $\mathbb{X} \to \mathbb{C}$, $p \mapsto \langle X_p, Y_p \rangle_p$ is $C^\infty(\mathbb{X})$. A manifold with a given Riemannian metric is called a *Riemannian manifold*. Let $g_{i,j} = \langle \partial_{i,p}, \partial_{j,p} \rangle_p$ and denote by g the matrix $(g_{i,j})$. The entries of g^{-1} are denoted be $g^{i,j}$. The Riemannian metric on $\mathbb{X}$ allows one to define a notion of length of a curve segment as well as the volume element (Riemannian measure) on $\mathbb{X}$. Also, if X is a vector field on $\mathbb{X}$, we may define $\|X\|_p := \langle X(p), X(p) \rangle_p$. The length of a differentiable curve $\gamma : [0, 1] \to \mathbb{X}$ is defined as $\mathcal{L}(\gamma) = \int_0^1 \|\gamma'(t)\|_{\gamma(t)} dt$. A differentiable curve $\gamma : [0, 1] \to \mathbb{X}$, such that the length of γ does not exceed that of any other piecewise differentiable curve joining $\gamma(0)$ to $\gamma(1)$, is called a *geodesic* [9, Proposition 3.6, Corollary 3.9]. For any $p \in \mathbb{X}$, there exists a neighborhood V of p, a number $\epsilon > 0$ and a C^∞ mapping $\gamma : (-2, 2) \times \mathcal{U} \to \mathbb{X}$, where $\mathcal{U} = \{(q, w) \in T\mathbb{X}; q \in V, w \in T_q\mathbb{X}, \|w\| < \epsilon\}$, such that $\gamma(\circ, q, w)$ is the unique geodesic of $\mathbb{X}$ with $\gamma(0, q, w) = q$, and the tangent vector at q being w [9, Proposition 2.7]. The gradient of a function $f \in C^\infty(\mathbb{X})$ is a vector field defined by

$$\nabla f = \sum_{i=1}^{n} \sum_{j=1}^{n} g^{i,j} \partial_i f \, \partial_j .$$

For the gradient field we have $\langle (\nabla f)_p, X_p \rangle_p = Xf(p)$ for every vector field X. A vector field F is *conservative* if for every $f \in C^\infty$ and curve $\gamma : [0, 1] \to \mathbb{X}$, $\int_0^1 (F(\gamma(t))f)(\gamma(t)) \|\gamma'(t)\|_{\gamma(t)} dt = f(\gamma(1)) - f(\gamma(0))$. The Laplace–Beltrami operator $\Delta^* f(p)$ is defined as the differential operator given by

$$\Delta^* f = -\frac{1}{\sqrt{|g|}} \sum_j \sum_{k=1}^{n} \partial_j (\sqrt{|g|} g^{j,k} \partial_k f),$$

where $|g| = \det(g)$. The operator Δ^* is an elliptic operator, so that the existence of a discrete spectrum and system of orthonormal eigenfunctions follows from the general theory of partial differential equations [36, Chap. 5.1]. The heat equation on $\mathbb{X}$ is the equation

$$\Delta^* U = \frac{\partial U}{\partial t}.$$

The solution of this equation subject to the condition $U(p,0)=f(p)$ is given by $\int K_t(p,q)f(p)d\mu(q)$, where μ is the Riemannian measure on $\mathbb{X}$ and K_t is the heat kernel described in our paper.

References

1. Belkin, M., Niyogi, P.: Semi-supervised learning on Riemannian manifolds. Mach. Learn. **56**, 209–239 (2004)
2. Belkin, M., Matveeva, I., Niyogi, P.: Regularization and regression on large graphs. In: Proc. of Computational Learning Theory, Banff, Canada (2004)
3. Boothby, W.M.: An Introduction to Differentiable Manifolds and Riemannian Geometry. Academic Press, San Diego (2003)
4. Cheng, S.Y., Li, P., Yau, S.-T.: On the upper estimate of the heat kernel of a complete Riemannian manifold. Am. J. Math. **103**(5), 1021–1063 (1981)
5. Chui, C.K., Donoho, D.L.: Special Issue: Diffusion maps and wavelets. Appl. Comput. Harmon. Anal. **21**(1) (2006)
6. Coifman, R.R., Maggioni, M.: Diffusion wavelets. Appl. Comput. Harmon. Anal. **21**, 53–94 (2006)
7. Davies, E.B.: L^p spectral theory of higher-order elliptic differential operators. Bull. Lond. Math. Soc. **29**, 513–546 (1997)
8. do Carmo, M.P.: Differential Geometry of Curves and Surfaces. Prentice Hall, New York (1976)
9. do Carmo, M.P.: Riemannian Geometry. Birkhäuser, Boston (1992)
10. Donoho, D.L., Grimes, C.: Image manifolds which are isometric to Euclidean space. http://www-stat.stanford.edu/~donoho/Reports/2002/WhenIsometry.pdf
11. Donoho, D.L., Levi, O., Starck, J.-L., Martinez, V.J.: Multiscale geometric analysis for 3-D catalogues. http://www-stat.stanford.edu/~donoho/Reports/2002/MGA3D.pdf
12. Dungey, N.: Some remarks on gradient estimates for heat kernels. Abstr. Appl. Anal. 2006, Art. ID 73020, 10 pp.
13. Filbir, F., Themistoclakis, W.: Polynomial approximation on the sphere using scattered data. Math. Nachr. **281**(5), 650–668 (2008)
14. Grigor'yan, A.: Upper bounds of derivatives of the heat kernel on an arbitrary complete manifold. J. Funct. Anal. **127**(2), 363–389 (1995)
15. Grigor'yan, A.: Estimates of heat kernels on Riemannian manifolds. In: Davies, E.B., Safarov, Y. (eds.) Spectral Theory and Geometry, Edinburgh, 1998. London Math. Soc. Lecture Note Ser., vol. 273, pp. 140–225. Cambridge University Press, Cambridge (1999)
16. Grigor'yan, A.: Heat kernels and function theory on metric measure spaces. Contemp. Math. **338**, 143–172 (2003)
17. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.-J.: Face recognition using laplacianfaces. IEEE Trans. Pattern Anal. Mach. Intell. **27**(3), 328–340 (2005)
18. Jones, P.W., Maggioni, M., Schul, R.: Universal local parametrizations via heat kernels and eigenfunctions of the Laplacian. Manuscript, arXiv:0709.1975
19. Keiner, J., Kunis, S., Potts, D.: Efficient reconstruction of functions on the sphere from scattered data. J. Fourier Anal. Appl. **13**, 435–458 (2007)
20. Kordyukov, Yu.A.: Lp-theory of elliptic differential operators on manifolds of bounded geometry. Acta Appl. Math. **23**(3), 223–260 (1991)
21. Korevaar, J.: Tauberian theory: A century of developments. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 329. Springer, Berlin (2004)
22. Korevaar, J.: A Tauberian boundedness theorem of the Ikehara type. Manuscript
23. Le Gia, Q.T., Mhaskar, H.N.: Localized linear polynomial operators and quadrature formulas on the sphere. SIAM J. Numer. Anal. **47**(1), 440–466 (2008)
24. Maggioni, M., Mhaskar, H.N.: Diffusion polynomial frames on metric measure spaces. Appl. Comput. Harmon. Anal. **24**(3), 329–353 (2008)
25. Mhaskar, H.N.: Approximation theory and neural networks. In: Jain, P.K., Krishnan, M., Mhaskar, H.N., Prestin, J., Singh, D. (eds.) Wavelet Analysis and Applications, Proceedings of the International Workshop in Delhi, 1999, pp. 247–289. Narosa Publishing, New Delhi (2001)
26. Mhaskar, H.N.: Polynomial operators and local smoothness classes on the unit interval, II. Jaén J. Approx. **1**(1), 1–25 (2009)

27. Mhaskar, H.N., Prestin, J.: Polynomial frames: a fast tour. In: Chui, C.K., Neamtu, M., Schumaker, L. (eds.) Approximation Theory XI, Gatlinburg, 2004, pp. 287–318. Nashboro Press, Brentwood (2005)
28. Mhaskar, H.N., Narcowich, F.J., Ward, J.D.: Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature. Math. Comput. **70**(235), 1113–1130 (2001). Corrigendum: Math. Comput. **71**, 453–454 (2001)
29. Minakshisundaram, S.: Eigenfunctions on Riemannian manifolds. J. Indian Math. Soc. (N.S.) **17**, 159–165 (1953–1954)
30. Nikolskii, S.M.: Approximation of Functions of Several Variables and Imbedding Theorems. Springer, New York (1975)
31. Oneill, B.: Elementary Differential Geometry. Elsevier, Amsterdam (2006)
32. Pesenson, I.: Bernstein-Nikolskii inequalities and Riesz interpolation formula on compact homogeneous manifolds. J. Approx. Theory **150**(2), 175–198 (2008)
33. Rivlin, T.J.: Chebyshev Polynomials. Wiley, New York (1990)
34. Schmid, D.: Scattered data approximation on the rotation group and generalizations. Ph.D. thesis, Technische Universität München (2009)
35. Sikora, A.: Riesz transform, Gaussian bounds and the method of wave equation. Math. Z. **247**, 643–662 (2004)
36. Taylor, M.E.: Partial Differential Equations, Basic Theory. Springer, New York (1996)
37. Xu, B.: Derivatives of the spectral function and Sobolev norms of eigenfunctions on a closed Riemannian manifold. Ann. Glob. Anal. Geom. **26**, 231–252 (2004)