Journal of Computational Physics 249 (2013) 96–112

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Minimum Sobolev norm interpolation with trigonometric polynomials on the torus

S. Chandrasekaran [a,1], K.R. Jayaraman [a,2], H.N. Mhaskar [b,c,*,3]

[a] Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, CA 93106, USA
[b] Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA
[c] Institute of Mathematical Sciences, Claremont Graduate University, Claremont, CA 91711, USA

ARTICLE INFO

Article history:
Received 20 September 2011
Received in revised form 29 September 2012
Accepted 19 March 2013
Available online 24 April 2013

Keywords:
Multivariate interpolation
Uniformly bounded interpolatory operators
Minimum Sobolev norm interpolation
Localized kernels

ABSTRACT

Let $q \geqslant 1$ be an integer, $\mathbf{y}_1, \ldots, \mathbf{y}_M \in [-\pi, \pi]^q$, and η be the minimal separation among these points. Given the samples $\{f(\mathbf{y}_j)\}_{j=1}^M$ of a smooth target function f of q variables, 2π-periodic in each variable, we consider the problem of constructing a q-variate trigonometric polynomial of spherical degree $\mathcal{O}(\eta^{-1})$ which interpolates the given data, remains bounded in the Sobolev norm (independent of η or M) on $[-\pi, \pi]^q$, and converges to the function f on the set where the data becomes dense. We prove that the solution of an appropriate optimization problem leads to such an interpolant. Numerical examples are given to demonstrate that this procedure overcomes the Runge phenomenon when interpolation at equidistant nodes on $[-1, 1]$ is constructed, and also provides a respectable approximation for bivariate grid data, which does not become dense on the whole domain.

1. Introduction

For clarity of exposition, the notation used in this and the next two sections may not be the same as in the rest of the paper.

In many engineering applications, one has to find a good approximation to an unknown *multivariate* target function which also interpolates the function at certain points, sometimes called landmarks. For example, in the problem of image registration, we are given a set of locations $x_j \in [-1, 1]^2$ in the first image and a corresponding set of points $y_j \in [-1, 1]^2$ in the second image. The idea is that the location x_j in the first image is the "same" as the location y_j in the second image. We then hope to find a map $g : [-1, 1]^2 \to \mathbb{R}^2$ such that $g(x_j) = y_j$, and such that g satisfies some smoothness conditions. There are at least two reasons for insisting on interpolatory approximation in this situation. First, the locations might have been chosen at great costs, including human efforts. Second, if the registration is being done many times over a sequence of images (for example when we stitch together video frames to form a large image), then a non-interpolatory approximation will cause a drift between the first image and the last image in the sequence.

In the univariate setting, perhaps the most classical method to obtain an interpolatory approximation is to use a polynomial interpolant. In the multivariate setting, it is not always possible to find a space of multivariate polynomials with dimension equal to the number of data points that admits an interpolant. For this reason, radial basis functions (RBF's) have

* Corresponding author address at: Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA. Tel.: +1 805 893 7542.
E-mail address: hmhaska@gmail.com (H.N. Mhaskar).

[1] The research of this author was supported, in part, by Grants CCF-0515320 and CCF-0830604 from the NSF.
[2] The research of this author was supported by Grant CCF-0830604 from the NSF.
[3] The research of this author was supported, in part, by Grant DMS-0908037 from the National Science Foundation and Grant W911NF-09-1-0465 from the U.S. Army Research Office.

0021-9991/$ - see front matter
http://dx.doi.org/10.1016/j.jcp.2013.03.041

become popular tools in recent years for multivariate interpolation problems. However, if one wants to use the RBF interpolant that minimizes a Sobolev norm, then there may not be an explicit closed form formula for this RBF kernel. In the periodic setting, it is easy to write down such a kernel in terms of a Fourier series (cf. (1.2)), which converges slowly. Therefore, the computation of such an RBF requires a careful approximation.

In this paper, we wish to explore a construction of multivariate polynomials directly to interpolate the values of a target function at arbitrarily chosen sites on a cube of the form $[-1,1]^q$, where $q>1$ is an integer. With a standard correspondence, this problem is the same as that of constructing multivariate trigonometric polynomials to interpolate the values of a target function at arbitrarily chosen sites $\mathbf{y}_1,\ldots,\mathbf{y}_M\in[-\pi,\pi]^q$. A very general result in [12] implies in particular that if

$$\eta=\min_{j\neq k,\,1\leqslant j,k\leqslant M}|\mathbf{y}_j-\mathbf{y}_k|,$$

then for every continuous function f on $[-\pi,\pi]^q$, 2π-periodic in each of its variables, there exists a trigonometric polynomial of order not exceeding a constant times η^{-1} which interpolates f at $\mathbf{y}_1,\ldots,\mathbf{y}_M$, and whose uniform norm does not exceed a constant times that of f, where the constants involved depend only upon q. However, the proof in [12] is not constructive.

In this paper, we are interested in interpolating functions in a Sobolev class W_s^p from the class $\mathbb{H}_N^q$ of trigonometric polynomials of spherical order N (see Section 4 for definitions). Since we must allow the degree N to be larger than the minimum required for obtaining an interpolant, the problem seems straightforward in the case when $p=2$. Let $M\geqslant 1$ be an integer, $\{\mathbf{y}_1,\ldots,\mathbf{y}_M\}\subset[-\pi,\pi]^q$, $f:[-\pi,\pi]^q\to\mathbb{C}$. The question of finding $T\in\mathbb{H}_n^q$ satisfying $T(\mathbf{y}_j)=f(\mathbf{y}_j)$ can be described in matrix notations as follows. Let D_N be the dimension of $\mathbb{H}_N^q$, A be a $M\times D_N$ matrix with $A_{j,\mathbf{k}}=\exp(i\mathbf{k}\cdot\mathbf{y}_j)$, and $\mathbf{f}$ be a column vector with j-th component given by $f(\mathbf{y}_j)$, $j=1,\ldots,M$. If $\mathbf{b}$ is a solution of

$$A\mathbf{b}=\mathbf{f}, \tag{1.1}$$

then $T(\mathbf{x})=\sum_{\mathbf{k}}b_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{x})$ solves the interpolation problem. It is rudimentary linear algebra to verify that the system (1.1) has a solution for every choice of $\mathbf{f}$ if and only if the rank of A is M; i.e., the Hermitian transpose A^* is one to one. In general, the solution will not be unique. A simple strategy to solve this system of equations is to minimize a weighted norm of $\mathbf{b}$. For example, minimizing $\sum_{\mathbf{k}}(|\mathbf{k}|^2+1)^s|b_{\mathbf{k}}|^2$ would lead to a solution with the minimal Sobolev norm in the sense of W_s^2. Naturally, we will refer to a solution obtained with this strategy (with possibly different variants such as considering $\sum_{\mathbf{k}}|\mathbf{k}|^{2s}|b_{\mathbf{k}}|^2$, or considering trigonometric polynomials of coordinatewise degree N, or considering only even trigonometric polynomials, etc.) as a *minimal Sobolev norm* (MSN) interpolant.

We remark that if we are not interested in interpolation with trigonometric polynomials, then the problem is well studied in the case of functions in $W_s^2, s>q/2$. In this case, the function K_{2s} defined by

$$K_{2s}(\mathbf{x})=\sum_{\mathbf{k}\in\mathbb{Z}^q}(|\mathbf{k}|^2+1)^{-s}\exp(i\mathbf{k}\cdot\mathbf{x}) \tag{1.2}$$

is continuous. The Golomb–Weinberger variation principle [6] can be used to show that the solution of the minimization problem

$$\text{minimize }\{\|g\|_{W_s^2}:\ g(\mathbf{y}_j)=f(\mathbf{y}_j),\ j=1,\ldots,M\} \tag{1.3}$$

has a solution in the span of $\{K_{2s}(\circ-\mathbf{y}_j)\}_{j=1}^M$, and therefore, can be found by solving an appropriate system of linear equations. However, when s is not a half integer, there is no closed form, easy to evaluate expression for K_{2s}. The series defining K_{2s} converges too slowly to allow an effective and easy computation. Therefore, we prefer to find an MSN interpolant as described above directly from the class $\mathbb{H}_N^q$.

Our objectives in this paper are the following:

- We prove in Theorem 5.1 that the MSN interpolant from the class $\mathbb{H}_N^q$ exists with a bounded W_s^p norm provided N is at least a constant times η^{-1}, where the constant may depend only on q,p, and s.
- In Theorem 5.2 we describe an explicit minimization problem to solve the interpolation problem. In particular, the minimization problems in the cases $p=1,\infty$ can be solved using linear programming (when the function is real-valued), and the case $p=2$ can be solved by solving an under-determined system of linear algebraic equation by standard numerical linear algebra methods.
- In Theorem 5.3, we prove that as the data set becomes dense on a portion of the torus, the MSN interpolants converge to the target function on this part, while their Sobolev norm remains bounded on the *entire* torus.

We postpone the discussion of the rates of convergence of the MSN interpolants to another paper in preparation.

The motivation behind this paper is both theoretical and numerical. In Section 2, we provide a motivation for our theoretical ideas in the classical univariate setting of interpolation by algebraic and trigonometric polynomials. In Section 3, we provide some numerical "proof of concept" experiments to illustrate the MSN algorithm in the case when $p=2$. The remainder of the paper is devoted to developing the theory. The notations used throughout this paper are introduced in Section 4. The main theorems are described in Section 5. The proofs of these results are given in Section 7. These require a great deal of preparation first. These preparatory results are developed in Section 6. This section may be skipped in a first reading, referring to the results as needed.

2. Background in the univariate setting

For each integer $n \geqslant 1$, let Π_n be the class of all (algebraic) polynomials of degree at most n, $\{y_{j,n}\}_{j=1}^n$ be a set of distinct points in $[-1,1]$, and $f:[-1,1]\to\mathbb{R}$. It is customary to organize the points $\{y_{j,n}\}_{j=1}^n$ as the nth row of an infinite matrix, known as the interpolation matrix Y. It is well known that there exists a unique polynomial $I_n(Y;f)\in\Pi_{n-1}$ such that $I_n(Y;f,y_{j,n}) = f(y_{j,n})$ for $j=1,\ldots,n$. The operator $I_n(Y)$ is a linear operator on the space of all the real valued functions on $[-1,1]$.

In the case when for each integer $n \geqslant 1$, the nth row of Y consists of n equidistant points on $[-1,1]$, the well known Runge example shows that for the function $f:[-1,1]\to\mathbb{R}$, $f(x) = (1+25x^2)^{-1}$, the sequence $\{I_n(Y;f)\}_{n=1}^\infty$ does not converge uniformly on $[-1,1]$. In this case, one can obtain a convergent sequence of interpolants by taking as the nth row of Y to the be zeros of the Chebyshev polynomial of degree n. However, Faber's theorem [14, Theorem 2, p. 27] states that for any interpolation matrix on $[-1,1]$, there exists a continuous function f on $[-1,1]$ such that the sequence $\{I_n(Y;f)\}$ does not converge uniformly.

The situation changes drastically if one allows the degree of the interpolatory polynomial to be greater than the minimal required for interpolation. Thus, the following Theorem 2.1 is a simple consequence of [16, Theorem 2.7, p. 52]. For $f:[-1,1]\to\mathbb{R}$, we write

$$\|f\|_{\infty,[-1,1]} := \sup_{t\in[-1,1]}|f(t)|, \quad \|f\|_{2,[-1,1]} := \left\{\frac{1}{\pi}\int_{-1}^{1}|f(t)|^2\frac{dt}{\sqrt{1-t^2}}\right\}^{1/2}.$$

We note that for n equidistant nodes on $[-1,1]$, the quantity d_n in the following theorem satisfies $d_n \geqslant 2/n$.

Theorem 2.1. *Let $y_{j,n} = \cos\theta_{j,n}\in[-1,1]$ be an arbitrary system of nodes ($0\leqslant\theta_{1,n}<\cdots<\theta_{n,n}\leqslant\pi$) and let*

$$d_n := \min_{1\leqslant j\leqslant n-1}(\theta_{j+1,n}-\theta_{j,n}).$$

Then for any $\epsilon>0$, there exist linear polynomial operators P_n on $C[-1,1]$ with the following properties:

(a)*If $m = \lfloor\pi(1+\epsilon)/d_n\rfloor$ then $P_n(P) = P$ for all $P\in\Pi_m$,For $f\in C[-1,1]$,*
(b)*$P_n(f)\in\Pi_N$ where $N = \lceil(\pi/d_n+1)(1+3\epsilon)\rceil$,*
(c)*$P(f,y_{j,n}) = f(y_{j,n})$ for $j=1,\ldots,n$,*
(d)*There exists an absolute constant $c>0$ such that*

$$\|f-P_n(f)\|_{\infty,[-1,1]} \leqslant c\inf_{P\in\Pi_m}\|f-P\|_{\infty,[-1,1]}. \tag{2.1}$$

A remarkable aspect of this theorem is that the points $y_{j,n}$ may all be concentrated only on a subinterval of $[-1,1]$, for example, $[0,1]$. Nevertheless, the sequence $\{P_n(f)\}$ converges to f on the whole interval $[-1,1]$. If we have a continuously differentiable function f, and require convergence on the part of the interval where the data becomes dense as $n\to\infty$, then a very simple construction can be given.

We recall that there is a one to one correspondence between functions on $[-1,1]$ and even, 2π-periodic function on $\mathbb{R}$, given by $f^\circ(\theta) = f(\cos\theta)$, $f:[-1,1]\to\mathbb{C}$. Moreover,

$$\|f^\circ\|_{\infty,[-\pi,\pi]} := \sup_{\theta\in[-\pi,\pi]}|f^\circ(\theta)| = \|f\|_{\infty,[-1,1]}, \quad \|f^\circ\|_{2,[-\pi,\pi]}^2 := \frac{1}{2\pi}\int_{-\pi}^{\pi}|f^\circ(\theta)|^2 d\theta = \|f\|_{2,[-1,1]}^2.$$

Proposition 2.1. *Let f° be continuously differentiable on $[-\pi,\pi]$. The minimization problem*

$$\textit{minimize } \|(P^\circ)'\|_{2,[-\pi,\pi]} \quad \textit{over all} \quad P\in\Pi_N, \quad \textit{subject to the constraints} \quad P(y_{j,n}) = f(y_{j,n}) \quad j=1,\ldots,n \tag{2.2}$$

has a solution P_n^* satisfying the following property. If Λ is a subsequence of positive integers, and x_0 is a limit point of a subsequence $\{y_{j,n}\}_{n\in\Lambda}$, then

$$\lim_{n\in\Lambda,n\to\infty}P_n^*(x_0) = f(x_0). \tag{2.3}$$

Proof. In this discussion, we will write P_n in place of $P_n(f)$, whose existence is asserted as in Theorem 2.1. In view of a theorem of Czipser and Freud [5], the estimate (2.1) implies that there exists an absolute constant $c_1>0$ such that

$$\|(P_n^\circ)'\|_{2,[-\pi,\pi]} \leqslant \|(P_n^\circ)'\|_{\infty,[-\pi,\pi]} \leqslant c_1\|(f^\circ)'\|_{\infty,[-\pi,\pi]}.$$

Thus, the minimization problem (2.2) has a feasible solution, P_n^* satisfying the same estimate as above. Further, for $\theta, \phi \in [-\pi, \pi]$,

$$|(P_n^*)^\circ(\theta) - (P_n^*)^\circ(\phi)| \leqslant \sqrt{2\pi|\theta-\phi|}\left\|\left((P_n^*)^\circ\right)'\right\|_{2,[-\pi,\pi]} \leqslant c_1\sqrt{|\theta-\phi|}\|(f^\circ)'\|_{\infty,[-\pi,\pi]}.$$

Since $P_n^*(y_{j,n}) = f(y_{j,n}),\ j = 1,\ldots,n$, the sequence $\{P_n^*\}$ is uniformly bounded and equicontinuous. In view of the Arzela–Ascoli theorem, this implies that any subsequence of the sequence $\{P_n^*\}$ has a uniformly convergent subsequence. If x_0 is a limit point of a subsequence $\{y_{j,n}\}_{n\in\Lambda}$, then it is not difficult to deduce using the interpolatory conditions that (2.3) holds. $\square$

In some situations it is also possible to construct approximating polynomials of degree lower than n (see [1]).

3. Numerical experiments

In this section, we will present numerical experiments that demonstrate the behavior of the MSN interpolant under different circumstances. In our computations below, we consider interpolation with algebraic polynomials of a suitable coordinate-wise degree. The standard substitution establishes a one to one correspondence between such polynomials and even trigonometric polynomials of the same order as the degree of the polynomials.

We first consider the classical Runge phenomenon by interpolating the function $f(x) = (1+25x^2)^{-1}$ at n equi-spaced points x_j in the interval $[-1,1]$. In Table 1 we show the results of our numerical experiments. The first column of the table shows the number of data points n used for interpolation. The remaining columns provide the maximum error in reconstructing the interpolant at 1024 equispaced samples in $[-1,1]$. The first column provides the error corresponding to a Wendland RBF. Thus, with

$$\omega(r) = (4r+1)*(1-r)_+^4, \quad r \geqslant 0,$$

we consider interpolation by sums of the form $\sum_j a_j\omega(|x-x_j|/\sigma)$. The parameter σ is chosen from the set $\{10, 30, 60, \ldots, 180\}$ to minimize the error in interpolating the given function on a 35×35 regular grid in $[-1,1]^2$. This range was chosen after a coarser choice of σ over a larger range of numbers was first used. The second column provides the maximum interpolation error obtained with MSN interpolation. The order of the MSN interpolant is

$$m = \left[\frac{6}{\min_{i\neq j}\left(\cos^{-1}(x_i) - \cos^{-1}(x_j)\right)}\right]. \tag{3.1}$$

The Sobolev parameter s is picked to be the largest possible value without suffering from numerical losses. This is essential to keep the order of convergence at the maximum. The MSN interpolant is computed using a weighted least-squares formulation as explained in another paper [4]. The last column provides the error in interpolating the considered Runge function with Matlab's cubic spline toolbox.

Next, we consider a two dimensional interpolation problem on a region inside the square $[-1.0, 1.0]\times[-1.0, 1.0]$. The function being reconstructed (shown in Fig. 1) is given by

$$f(x,y) = \frac{1}{1+25(x^2+y-0.3)^2} + \frac{1}{1+25(x+y-0.4)^2} + \frac{1}{1+25(x+y^2-0.5)^2} + \frac{1}{1+25(x^2+y^2-0.25)^2}.$$

The function has the Runge structure along two parabolii, a circle and a straight line. Table 2 shows the maximum error in reconstructing $f(x,y)$ at a $(n+11)\times(n+11)$ grid using the MSN interpolant at an $n\times n$ grid. The error is normalized by the maximum value of $f(x,y)$ over the reconstruction points. The parameter s is varied in steps of 2 from 2 to 12. The degree of the interpolating polynomial was chosen analogously to (3.1). Table 2 also shows the accuracy obtained with the Wendland RBF interpolant, in its last column. Compared to the RBF and Spline interpolants, the MSN interpolant has an order or more in its accuracy for sufficiently large s and increasing n.

These experiments show that the proposed scheme can perform well on difficult problems in two dimensions where traditional interpolation schemes require much more work to achieve comparable accuracy.

The proposed method relies on a complete orthogonal decomposition (CODA) technique as described in [4] for numerical stability. For the sake of completeness we give a brief overview of the numerical technique. The numerical solution of the MSN interpolation problem when $p = 2$ boils down to the following optimization problem

Table 1
Interpolation of 1D Runge function. Column 2: interpolation error with Wendland RBF. Column 3: interpolation error with MSN, $s = 6$. Column 4: interpolation error with Cubic Splines.

n	errRBF	errMSN	errSpline
10	1.4e−01	1.4e−01	1.4e−01
100	6.6e−06	2.8e−11	6.9e−06
200	2.2e−06	2.3e−13	3.7e−07
300	1.4e−06	2.2e−14	7.7e−08

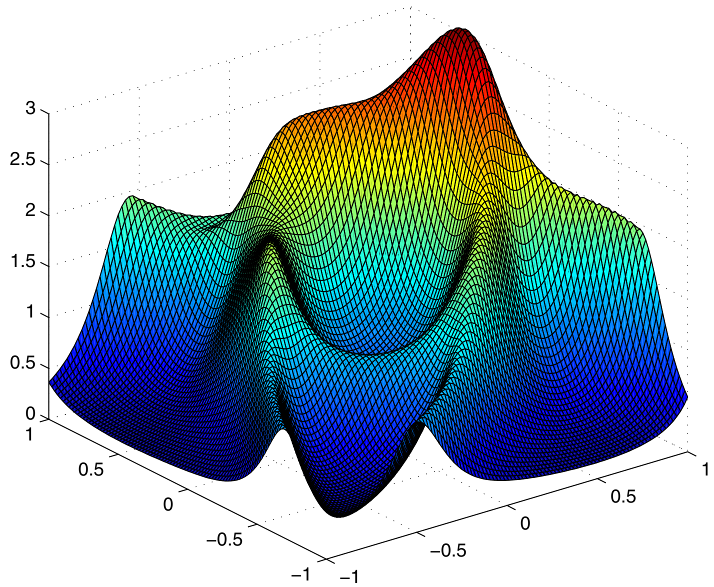

Fig. 1. The test function $f(x,y)$.

Table 2
Maximum error of MSN interpolant in region $[-1.0, 1.0]^2$.

n	$s=2$	$s=4$	$s=6$	$s=8$	$s=10$	$s=12$	RBF	Spline
10	1.1e−01	1.2e−01	2.3e−01	4.9e−01	1.0e+00	1.6e+00	1.5e−01	1.5e−01
20	3.1e−02	2.7e−02	3.6e−02	1.2e−01	5.2e−01	2.7e+00	4.1e−02	4.5e−02
40	6.7e−03	2.2e−03	1.4e−03	1.4e−03	4.3e−03	1.9e−02	7.6e−04	3.5e−03
80	9.5e−04	9.4e−05	1.7e−05	4.6e−06	4.3e−06	5.7e−06	4.1e−05	1.5e−04
100	5.1e−04	3.8e−05	6.2e−06	1.5e−06	5.0e−07	4.2e−07	7.2e−05	5.0e−05

$$\min_{Vb=f} \|Db\|_2,$$

where b is the vector of unknown Chebyshev coefficients, V is the Chebyshev–Vandermonde matrix, and D is a diagonal positive-definite matrix with condition number $O(n^s)$. Intuitively speaking, the matrix V is made fat enough that it's condition number can be viewed to be a constant. So the condition number of the problem is essentially that of D. Such scaled minimum 2-norm problems can be solved by utilizing ideas suggested in [7] or [2]. Both these references are concerned with diagonally weighted least-squares problems, but their approach can be readily adopted to the diagonally weighted minimum 2-norm problem. The last reference [2] in particular also has a stability analysis. More details can be found in [4]

The ideas presented here can also be generalized to handle noisy and redundant observations. These matters are reported elsewhere [3].

4. Notations

In the sequel, $q \geqslant 1$ will denote a fixed integer, and we will think of 2π-periodic functions on $\mathbb{R}^q$ as functions on $[-\pi,\pi]^q$, tacitly identified with the q dimensional torus. Analogous to the univariate case, any function $f:[-1,1]^q \to \mathbb{C}$, corresponds uniquely to the 2π–periodic function f° on $\mathbb{R}^q$ by the *standard correspondence*

$$f^\circ(\theta_1,\ldots,\theta_q) = f(\cos\theta_1,\ldots,\cos\theta_q).$$

If $\mathbf{x} \in \mathbb{R}^q$, we write

$$|\mathbf{x}|_p := \begin{cases} \left\{ \sum_{k=1}^{q} |x_k|^p \right\}^{1/p}, & \text{if} \quad 1 \leqslant p < \infty, \\ \max_{1\leqslant k\leqslant q} |x_k|, & \text{if} \quad p = \infty. \end{cases}$$

For brevity, we will use the same notation for vectors with different dimensions. If $p = 2$ then we write $|\cdot|$ to denote $|\cdot|_2$. Although this does not constitute any formal conflict of notations, we hope that it will be clear from the context whether the notation means the ℓ^2 norm of a vector or the absolute value of a real number.

Let $\mathbb{H}_n^q$ denote the class of all trigonometric polynomials in q variables with spherical order (or degree) at most n; i.e.,

$$\mathbb{H}_n^q := \left\{ \sum_{\mathbf{k}\in\mathbb{Z}^s, |\mathbf{k}|\leqslant n} a_{\mathbf{k}} \exp(i\mathbf{k}\cdot(\circ)) : a_{\mathbf{k}} \in \mathbb{C} \right\}.$$

Here, we find it convenient to use the same notation even if n is not an integer. We note that our theory will work also if we consider coordinatewise or total degree in place of spherical degree. However, since the Sobolev classes are defined in terms of $|\mathbf{k}|$, it is convenient in theoretical considerations to work with polynomials with different spherical degrees. It is not difficult to see that in the standard correspondence, multivariate algebraic polynomials on $[-1,1]^q$ correspond to the trigonometric polynomials of the same order which are symmetric in each of the variables. Therefore, in this paper, we focus mainly in the interpolation of multivariate periodic functions. The results can also be applied trivially to the interpolation of functions on $[-1,1]^q$, with suitable smoothness conditions defined in terms of the corresponding periodic function.

If $1 \leqslant p \leqslant \infty, K \subset [-\pi,\pi]^q$ and $f : K \to \mathbb{C}$ are Lebesgue measurable, we write

$$\|f\|_{p,K} = \begin{cases} \left\{\int_K |f(\mathbf{x})|^p d\mathbf{x}\right\}^{1/p}, & \text{if} \quad 1 \leqslant p < \infty, \\ \operatorname{ess\,sup}_{\mathbf{x}\in K} |f(\mathbf{x})|, & \text{if} \quad p = \infty. \end{cases} \tag{4.1}$$

The symbol $L^p(K)$ denotes the class of all Lebesgue measurable functions f for which $\|f\|_{p,K} < \infty$, with the usual convention that two functions are considered equal if they are equal almost everywhere. If $K = [-\pi,\pi]^q$, we will omit its mention from the notations. If $1 < p < \infty$, we will write $p' := p/(p-1)$, and extend this notation to $p = 1, \infty$ by setting $1' = \infty$, $\infty' = 1$. Analogous notations will be used to denote the L^p norms of functions defined on other sets; for example, subsets of $\mathbb{R}$ or $\mathbb{R}^q$. We do not find it necessary to complicate our notations by using different notations to denote these nominally different objects.

If $f \in L^1$, the Fourier coefficients of f are defined by

$$\hat{f}(\mathbf{k}) := \frac{1}{(2\pi)^q} \int_{[-\pi,\pi]^q} f(\mathbf{x}) \exp(-i\mathbf{k}\cdot\mathbf{x}) d\mathbf{x}, \quad \mathbf{k} \in \mathbb{Z}^q. \tag{4.2}$$

If $f \in L^p$, then its degree of approximation from $\mathbb{H}_n^q$ is defined by

$$E_{n,p}(f) := \inf_{T\in\mathbb{H}_n^q} \|f - T\|_p.$$

If $s \in \mathbb{R}, 1 \leqslant p \leqslant \infty$, the Sobolev class W_s^p consists of all $f \in L^p$ for which there exists $f^{(s)} \in L^p$ such that

$$\widehat{f^{(s)}}(\mathbf{k}) = (|\mathbf{k}|^2 + 1)^{s/2} \hat{f}(\mathbf{k}), \quad \mathbf{k} \in \mathbb{Z}^q.$$

We define

$$\|f\|_{W_s^p} := \|f^{(s)}\|_p, \tag{4.3}$$

and note that W_s^p is a Banach space. We observe that if Δ is the Laplacian operator on $\mathbb{R}^q$, and s is an even, positive integer, then $f^{(s)} = (I - \Delta)^{s/2} f$, where I is the indentity operator. In particular, in this case, the operator $f \mapsto f^{(s)}$ is a surface derivative operator on the torus identified with $[-\pi,\pi]^q$.

4.1. The constant convention

The symbols $c, c_1, \ldots,$ will denote generic positive constants, depending on such fixed parameters of the problem as p, s, q, etc. and other quantities explicitly indicated, but their value may different at different occurrences, even within a single formula. The notation $A \sim B$ means that $c_1 A \leqslant B \leqslant c_2 A$.

The following proposition, to be proved in Section 6.2, gives an integral representation of functions in W_s^p.

Proposition 4.1. *Let $1 \leqslant p \leqslant \infty,\ s > q/p$. Then there exists a function $K_s \in L^{p'}$ such that*

$$\widehat{K_s}(\mathbf{k}) = (|\mathbf{k}|^2 + 1)^{-s/2}, \quad \mathbf{k} \in \mathbb{Z}^q. \tag{4.4}$$

If $f \in W_s^p$, then for almost all $\mathbf{x} \in [-\pi,\pi]^q$,

$$f(\mathbf{x}) = \frac{1}{(2\pi)^q} \int_{[-\pi,\pi]^q} K_s(\mathbf{x} - \mathbf{y}) f^{(s)}(\mathbf{y}) d\mathbf{y} = \frac{1}{(2\pi)^q} \int_{[-\pi,\pi]^q} K_s(\mathbf{y} - \mathbf{x}) f^{(s)}(\mathbf{y}) d\mathbf{y}. \tag{4.5}$$

In particular, f is almost everywhere equal to a continuous function. Denoting this continuous function again by f, we have for any integer $n \geqslant 1$,

$$\|f\|_\infty \leqslant c\|f\|_{W^p_s}, \quad E_{n,\infty}(f) \leqslant cn^{q/p-s}\|f\|_{W^p_s}. \tag{4.6}$$

As customary in the theory of interpolation, let Y be the *interpolation matrix* whose nth row Y_n contains M_n points $\{\mathbf{y}_{j,n}\}_{j=1}^{M_n} \subset [-\pi,\pi]^q$. We note that Y is not a matrix in the usual sense, it is only a visualization of our data, referred to as a matrix to draw an analogy with the univariate case. If $\mathcal{C} \subseteq [-\pi,\pi]^q$ we define the *separation radius* $\eta(\mathcal{C})$ of $\mathcal{C}$ by

$$\eta(\mathcal{C}) := (1/2)\inf_{\mathbf{x},\mathbf{y}\in\mathcal{C},\mathbf{x}\neq\mathbf{y}} |\mathbf{x}-\mathbf{y}|. \tag{4.7}$$

We will simplify our notation, and write $\eta_n := \eta(Y_n)$.

5. Main results

Our first theorem is an analogue of Theorem 2.1 for multivariate trigonometric polynomials and Sobolev norms rather than the supremum norm. We will comment about the proof of this theorem towards the end of this section.

Theorem 5.1. *Let $1 \leqslant p \leqslant \infty, s > q/p, Y$ be as described in Section 4. There exists an integer N^* with $N^* \sim \eta_n^{-1}$ and a mapping $\mathbf{P} : W^p_s \to \mathbb{H}^q_{N^*}$ such that for every $f \in W^p_s$,*

$$\mathbf{P}(f, \mathbf{y}_{j,n}) = f(\mathbf{y}_{j,n}), \quad j = 1, \ldots, M_n, \tag{5.1}$$

and

$$\|f - \mathbf{P}(f)\|_{W^p_s} \leqslant c\inf\{\|f - T\|_{W^p_s} : T \in \mathbb{H}^q_{N^*}\}. \tag{5.2}$$

Theorem 5.2. *We assume the set up as in Theorem 5.1. We consider the minimization problem*

$$\textit{minimize}\left\{\frac{1}{N^{*q}}\sum_{0\leqslant\mathbf{k}\leqslant 3N^*-1}|P^{(s)}(2\pi\mathbf{k}/(3N^*))|^p\right\}^{1/p}, \tag{5.3}$$

where the minimum is over all $P \in \mathbb{H}^q_{N^*}$, such that $P(\mathbf{y}_{j,n}) = f(\mathbf{y}_{j,n}),\ j = 1, \ldots, M_n$, and an appropriate interpretation is understood in the case $p = \infty$. There exists a solution of this problem, $\mathbb{T}^*_n = \mathbb{T}^*_n(p, Y_n, f) \in \mathbb{H}^q_{N^*}$, such that $\|\mathbb{T}^*_n\|_{W^p_s} \leqslant c\|f\|_{W^p_s}$.

We note that the problem (5.3) has a unique solution if $p = 2$, and in this case, the corresponding operator $\mathbb{T}^*_n$ is linear in f. The proof of Theorem 5.2 is a simple consequence of Theorem 5.1 and a discretization inequality proved in Lemma 7.1.

The next theorem examines the convergence properties of the sequence $\{\mathbb{T}^*_n\}$. If $K \subseteq [-\pi,\pi]^q$ and $x \in [-\pi,\pi]^q$, we define

$$\mathrm{dist}(K, x) := \inf_{y\in K}|x - y|_\infty.$$

Theorem 5.3. *Let $1 \leqslant p \leqslant \infty, s > q/p, f \in W^p_s,\ N^*$ and $\mathbb{T}^*_n$ be found as in Theorem 5.2. If Λ is a subsequence of positive integers, $\mathbf{x}_0 \in [\pi,\pi]^q$, and*

$$\lim_{\substack{n\to\infty\\ n\in\Lambda}} \mathrm{dist}(Y_n, \mathbf{x}_0) = 0, \tag{5.4}$$

then

$$\lim_{\substack{n\to\infty\\ n\in\Lambda}} \mathbb{T}^*_n(\mathbf{x}_0) = f(\mathbf{x}_0).$$

The proof of Theorem 5.3, as expected, is a compactness argument. We also need to estimate the discrete norm used in (5.3) by the corresponding continuous norm. The necessary facts are stated in Lemmas 7.1 and 7.2.

The proof of Theorem 5.1 occupies a major part of this paper. Using an idea from [12], we will first prove the following general theorem for the feasibility of matrix equations similar to (1.1).

Theorem 5.4. *Let $D \geqslant M \geqslant 1$ be integers, A be a $M \times D$ matrix with complex entries, A^* be its Hermitian conjugate. Let $|||\cdot|||_M$ (respectively $|||\cdot|||^*_D$) be norms on $\mathbb{C}^M$ (respectively, $\mathbb{C}^D$), and $|||\cdot|||^*_M$ (respectively $|||\cdot|||_D$) be the corresponding dual norms. If there exists $\kappa > 0$ such that*

$$|||\mathbf{c}|||^*_M \leqslant \kappa|||A^*\mathbf{c}|||^*_D, \quad \mathbf{c} \in \mathbb{C}^M, \tag{5.5}$$

then for every $\mathbf{f} \in \mathbb{C}^M$, there exists $\mathbf{b} \in \mathbb{C}^D$ such that

$$A\mathbf{b} = \mathbf{f}, \tag{5.6}$$

and

$$|||\mathbf{b}|||_D \leqslant \kappa|||\mathbf{f}|||_M. \tag{5.7}$$

In order to apply Theorem 5.4, we will choose a suitable integer N^*. With this choice, we choose $M = M_n,\ D$ to be the dimension of $\mathbb{H}_{N^*}$, the matrix A to be $(\exp(-i\mathbf{k}\cdot\mathbf{y}_{j,n}))_{|\mathbf{k}|\leqslant N^*,\ j=1,\ldots,M_n}$, the norm $|||\cdot|||_M$ to be the dual norm of the norm

$$|||\mathbf{c}|||_M^* := \left\|\sum_{j=1}^{M} c_j K_s(\circ - \mathbf{y}_j)\right\|_{p'}, \tag{5.8}$$

and

$$|||\mathbf{b}|||_D := \left\|\sum_{|\mathbf{k}|\leqslant N^*} (|\mathbf{k}|^2+1)^{s/2} b_{\mathbf{k}} \exp(i\mathbf{k}\cdot\circ)\right\|_p. \tag{5.9}$$

If $f \in W_s^p$, and $\mathbf{f} \in \mathbb{C}^M$ is the column vector with entries $f(\mathbf{y}_j)$, then using the kernel representation (4.5), it is not difficult to verify that

$$|||\mathbf{f}|||_M \leqslant \|f\|_{W_s^p}.$$

If (5.5) is satisfied then the conclusions (5.6) and (5.7) imply the existence of an interpolant in $\mathbb{H}_{N^*}$ with the W_s^p norm bounded by a constant multiple of $\|f\|_{W_s^p}$. The proof of Theorem 5.1 then becomes easy.

Thus, the proof of Theorem 5.1 depends upon proving (5.5) with the value of N^* as indicated in Theorem 5.1. As expected, this involves a careful approximation of a function of the form $G = \sum_{j=1}^{M} c_j K_s(\circ - \mathbf{y}_j)$ by elements of $\mathbb{H}_{N^*}$. This is given in Theorem 6.3. The proof of this theorem is an adaptation of the ideas in [9,11,10]. In particular, this involves an estimation of the coefficients of G in terms of the norm of G, as well as a good approximation bound on K_s.

6. Technical preparation

In this section, we present many technical results which are preparatory to the proof of the main results of Section 5. A main ingredient in our proof of Theorem 5.1 is Theorem 6.3. Sections 6.1 and 6.2 are devoted to the proof of this theorem. In Section 6.1, we introduce a localized kernel and the corresponding operator which will be used throughout this paper, and prove a number of results regarding these. In particular, we use these results in Section 6.2 to prove Proposition 4.1 and establish a few other facts related to the kernel K_s.

6.1. Localized kernels

Let $q \geqslant 1$ be an integer. For $t > 0$, and $H : \mathbb{R}^q \to \mathbb{R}$, we define formally

$$\Psi_t(H, \mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^q} H(\mathbf{k}/t)\exp(i\mathbf{k}\cdot\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^q. \tag{6.1}$$

We set $\Psi_0(H, \mathbf{x}) := H(0)$. If $S \geqslant 1$ is an integer, and H is S times continuously differentiable, we will use the notation

$$\mathcal{N}(H) := \mathcal{N}_S(H) := \max_{\mathbf{0}\leqslant\mathbf{k}\leqslant S} \|D^{\mathbf{k}}H\|_{1,\mathbb{R}^q}.$$

The following theorem summarizes the important localization estimate for the kernel Ψ_t.

Theorem 6.1. *Let $S > q$ be an integer, $b > 0$, and $H : \mathbb{R}^q \to \mathbb{R}$ be an S times continuously differentiable function such that $H(\mathbf{x}) = 0$ if $|\mathbf{x}| \geqslant b$.*

(a) *For $\mathbf{x} \neq 0,\ \mathbf{x} \in [-\pi, \pi]^q,\ bt \geqslant 1/2$,*

$$|\Psi_t(H; \mathbf{x})| \leqslant ct^q\mathcal{N}(H)\min\{b^q, (t|\mathbf{x}|)^{-S}\}, \tag{6.2}$$

and for $1 \leqslant p \leqslant \infty$,

$$\|\Psi_t(H, \circ)\|_p \leqslant ct^{q/p'}\mathcal{N}(H)(b^q)^{1-q/(pS)}. \tag{6.3}$$

(b) *Moreover, if*

$$H(\mathbf{x}) \geqslant \alpha\mathcal{N}(H) > 0, \quad \text{and} \quad a \leqslant |\mathbf{x}| \leqslant a_1, \tag{6.4}$$

for suitable positive numbers α, a, a_1, then for $t \geqslant 1/2$,

$$\max_{\mathbf{x}\in[-\pi,\pi]^q} |\Psi_t(H, \mathbf{x})| = \Psi_t(H, \mathbf{0}) \geqslant c_1(\alpha, a, a_1)t^q\mathcal{N}(H). \tag{6.5}$$

Here, the constants denoted by c may depend upon q and S only.

We observe that the variable $\mathbf{x}$ in the localization estimate (6.2) is limited to a compact set, so that for any $\delta > 0$, and $|\mathbf{x}| \geqslant \delta$, the interest in the estimate is the term t^{-S}. We note too that the quantity $(t|\mathbf{x}|)^{-S}$ is the dominant term in (6.2) if $|\mathbf{x}| > 1/(b^{q/S}t)$, where t is a potentially large parameter.

The proof of this theorem is essentially the same as that of the univariate version of this theorem proved in [13]. We reproduce it for the sake of completion.

For $F \in L^1(\mathbb{R}^q)$, we define its Fourier transform by

$$\hat{F}(\mathbf{x}) = \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q} F(\mathbf{y})\exp(-i\mathbf{y}\cdot\mathbf{x})d\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^q. \tag{6.6}$$

and the inverse Fourier transform by

$$\tilde{F}(\mathbf{x}) = \int_{\mathbb{R}^q} F(\mathbf{y})\exp(i\mathbf{y}\cdot\mathbf{x})d\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^q.$$

The following well known proposition lists the properties of Fourier transform which we will need in this section.

Proposition 6.1.

(a) *If both F and $\hat{F}$ are in $L^1(\mathbb{R}^q)$, then the Fourier inversion formula holds*:

$$F(\mathbf{x}) = \tilde{\hat{F}}(\mathbf{x}) = \hat{\tilde{F}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^q.$$

(b) *If $S \geqslant 1$ is an integer, F is compactly supported, and is S times continuously differentiable then*

$$|\hat{F}(\mathbf{x})| \leqslant \frac{c\mathcal{N}(F)}{|\mathbf{x}|^S}, \quad |\tilde{F}(\mathbf{x})| \leqslant \frac{c\mathcal{N}(F)}{|\mathbf{x}|^S}, \quad \mathbf{x} \in \mathbb{R}^q, \ \mathbf{x} \neq 0. \tag{6.7}$$

(c) *Let $F \in L^1(\mathbb{R}^q)$. Then, if $\mathbb{K} := [-\pi,\pi]^q$, we have*

$$\|F\|_{1,\mathbb{R}^q} = \sum_{\mathbf{k}\in\mathbb{Z}^q} \|F(\circ + 2\pi\mathbf{k})\|_{1,\mathbb{K}} < \infty.$$

In particular, the function

$$f(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^q} F(\mathbf{x} + 2\pi\mathbf{k})$$

is defined for almost all $\mathbf{x} \in [-\pi,\pi]^q$, $f \in L^1(\mathbb{K})$ and $\|f\|_{1,\mathbb{K}} = \|F\|_{1,\mathbb{R}^q}$. If F is $S > q$ times continuously differentiable, then we have the following Poisson summation formula valid for every $\mathbf{x} \in [-\pi,\pi]^q$:

$$\sum_{\mathbf{m}\in\mathbb{Z}^q} \hat{F}(\mathbf{m})\exp(i\mathbf{m}\cdot\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^q} F(\mathbf{x} + 2\pi\mathbf{k}), \quad \mathbf{x} \in [-\pi,\pi]^q. \tag{6.8}$$

Proof. Part (a) is given in [15, Chapter 1, Corollary 1.21]. Part (b) is a simple consequence of [15, Chapter 1, Theorem 1.1 and formula (1.9)]. Part (c) follows from part (b) and [15, Chapter VII, Theorem 2.4 and Corollary 2.6]. □

Proof of Theorem 6.1. In this proof, we may assume without loss of generality that $\mathcal{N}(H) = 1$.

Since H is continuous and compactly supported, $H \in L^1(\mathbb{R}^q)$. Also, the series in (6.1) is only a trigonometric polynomial, and in particular, absolutely and uniformly convergent. Further, since H is $S > q$ times continuously differentiable, we deduce from Proposition 6.1 that both the Fourier inversion formula and the Poisson summation formula hold for all $\mathbf{x} \in \mathbb{R}^q$. Since the Fourier transform of $H(\circ/t)$ is given by $t^q\hat{H}(t\circ)$, we obtain

$$\Psi_t(H;\mathbf{x}) = t^q \sum_{\mathbf{k}\in\mathbb{Z}^q} \tilde{H}(t(\mathbf{x} + 2\mathbf{k}\pi)), \qquad \mathbf{x} \in \mathbb{R}^q. \tag{6.9}$$

Let $bt \geqslant 1/2$. Since $H(\mathbf{k}/t) = 0$ if $|\mathbf{k}| \geqslant bt$, and $|\{\mathbf{k} \in \mathbb{Z}^q; |\mathbf{k}| < bt\}| \leqslant c(bt)^q$, it clear that

$$|\Psi_t(H;\mathbf{x})| \leqslant c(bt)^q\mathcal{N}(H). \tag{6.10}$$

Since H is S times continuously differentiable, we now use (6.9) and (6.7) to obtain for $\mathbf{x} \in [-\pi,\pi]^q, \mathbf{x} \neq 0, bt \geqslant 1/2$,

$$|\Psi_t(H;\mathbf{x})| \leqslant ct^q \sum_{\mathbf{k}\in\mathbb{Z}^q} \frac{1}{t^S|\mathbf{x} + 2\mathbf{k}\pi|^S} = c\frac{t^{q-S}}{|\mathbf{x}|^S} + ct^{q-S} \sum_{\mathbf{k}\in\mathbb{Z}^q, |\mathbf{k}|_\infty \geqslant 1} \frac{1}{|\mathbf{x} + 2\mathbf{k}\pi|^S}. \tag{6.11}$$

Since $|\mathbf{x}|_\infty \leqslant \pi$, we obtain for $\mathbf{k} \in \mathbb{Z}^q, |\mathbf{k}|_\infty \geqslant 1$, that

$$|\mathbf{x}+2\mathbf{k}\pi| \geqslant |\mathbf{x}+2\mathbf{k}\pi|_\infty \geqslant 2\pi|\mathbf{k}|_\infty - |\mathbf{x}|_\infty \geqslant |\mathbf{x}|_\infty(2|\mathbf{k}|_\infty - 1) \geqslant \frac{|\mathbf{x}|}{\sqrt{q}}(2|\mathbf{k}|_\infty - 1).$$

Therefore, for $\mathbf{x} \in [-\pi,\pi]^q,\ \mathbf{x} \neq 0$, we deduce that

$$\sum_{\substack{\mathbf{k} \in \mathbb{Z}^q \\ |\mathbf{k}|_\infty \geqslant 1}} \frac{1}{|\mathbf{x}+2\mathbf{k}\pi|^S} \leqslant \frac{c}{|\mathbf{x}|^S}\sum_{j=1}^{\infty}\sum_{|\mathbf{k}|_\infty = j}\frac{1}{(2|\mathbf{k}|_\infty - 1)^S} = \frac{c}{|\mathbf{x}|^S}\sum_{j=1}^{\infty}|\{\mathbf{k}\in\mathbb{Z}^q : |\mathbf{k}|_\infty = j\}|\frac{1}{(2j-1)^S}$$

Since $|\{\mathbf{k} : |\mathbf{k}|_\infty \leqslant j\} = (2j+1)^q$ for all $j \geqslant 0$, we have

$$\left|\{\mathbf{k}\in\mathbb{Z}^q : |\mathbf{k}|_\infty = j\}\right| = (2j+1)^q - (2j-1)^q = \sum_{\ell=1}^{q}\binom{q}{\ell}2^\ell(2j-1)^{q-\ell} \leqslant c(2j-1)^{q-1}.$$

We conclude that

$$\sum_{\substack{\mathbf{k} \in \mathbb{Z}^q \\ |\mathbf{k}|_\infty \geqslant 1}} \frac{1}{|\mathbf{x}+2\mathbf{k}\pi|^S} = \frac{c}{|\mathbf{x}|^S}\sum_{j=1}^{\infty}|\{\mathbf{k}\in\mathbb{Z}^q : |\mathbf{k}|_\infty = j\}|\frac{1}{(2j-1)^S} = \frac{c}{|\mathbf{x}|^S}\sum_{j=1}^{\infty}(2j-1)^{q-1}\frac{1}{(2j-1)^S}. \tag{6.12}$$

Since $S > q$, the infinite series in (6.12) converges. Therefore, (6.11)and (6.10) lead to (6.2).

To prove (6.3) with $1 \leqslant p < \infty$, let $r = b^{-q/S}$. Using (6.2) and our choice of r, we deduce that

$$\begin{aligned}\int_{[-\pi,\pi]^q}|\Psi_t(H;\mathbf{x})|^p d\mathbf{x} &= \int_{\{\mathbf{x}\in[-\pi,\pi]^q,\ |\mathbf{x}|\leqslant r/t\}}|\Psi_t(H;\mathbf{x})|^p d\mathbf{x} + \int_{\{\mathbf{x}\in[-\pi,\pi]^q,\ |\mathbf{x}|\geqslant r/t\}}|\Psi_t(H;\mathbf{x})|^p d\mathbf{x} \\ &\leqslant c_2\left\{(bt)^{qp}\int_{\{\mathbf{x}\in\mathbb{R}^q,\ |\mathbf{x}|\leqslant r/t\}}d\mathbf{x} + (t^{q-S})^p\int_{\{\mathbf{x}\in\mathbb{R}^q,\ |\mathbf{x}|\geqslant r/t\}}\frac{d\mathbf{x}}{|\mathbf{x}|^{Sp}}\right\} \leqslant c_3\left\{(bt)^{qp}(r/t)^q + t^{(q-S)p}(r/t)^{q-Sp}\right\}.\end{aligned}$$

In view of our choice of r, this leads to (6.3). This completes the proof of part (a).

If (6.4) holds, then (6.5) is an immediate consequence of the definitions. □

If $f \in L^1$, we define

$$\sigma_t(H,f,\mathbf{x}) := \frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} f(\mathbf{y})\Psi_t(H,\mathbf{x}-\mathbf{y})d\mathbf{y}. \tag{6.13}$$

An important consequence of (6.3) is the following.

Corollary 6.1. *Let H be as in Theorem 6.1(a). Then for $1 \leqslant p \leqslant \infty$, $f \in L^p$, we have*

$$\|\sigma_t(H,f)\|_p \leqslant c\mathcal{N}(H)\max(b^{1-q},(b^q)^{1-q/S})\|f\|_p, \quad t \geqslant 0. \tag{6.14}$$

Proof. If $0 \leqslant bt < 1/2, \sigma_t(H,f) = H(\mathbf{0})\hat{f}(\mathbf{0})$, and (6.14) is trivial. Let $bt \geqslant 1/2$. In view of (6.3), $\|\Psi_t(H,\circ)\|_1 \leqslant c\mathcal{N}(H)(b^q)^{1-q/S}$. The estimate (6.14) is now clear in the case $p = \infty$, and follows from Fubini's theorem in the case when $p = 1$. An application of Riesz–Thorin theorem leads to the intermediate cases. □

Our next major goal is to prove Theorem 6.2, which is required in the proof of Theorem 6.3. Let $\{\mathbf{y}_j\}_{j=1}^M \subset [-\pi,\pi]^q, m \geqslant 1$ be an integer with

$$\min_{j\neq k}|\mathbf{y}_j - \mathbf{y}_k| \geqslant 1/m. \tag{6.15}$$

We note that this implies $M \leqslant cm^q$. In the sequel, we will assume tacitly that $\{\mathbf{y}_j\}_{j=1}^M$ is one of the members of a sequence of finite subsets of $[-\pi,\pi]^q$. Thus, M and m are variables, and the constants are independent of these.

Theorem 6.2. *Let H be as in Theorem 6.1(a) and supported on $\{|\mathbf{x}| \leqslant 1\}$. Let $n \geqslant 1$ be an integer, $1 \leqslant p \leqslant \infty$, $\mathbf{a} \in \mathbb{R}^M$, and*

$$G_n(\mathbf{x}) := \sum_{j=1}^{M} a_j\Psi_n(H,\mathbf{x}-\mathbf{y}_j), \quad \mathbf{x} \in [-\pi,\pi]^q.$$

(a) *We have*

$$\|G_n\|_p \leqslant cn^{q/p'}\left\{1+(m/n)^S\right\}^{1/p'}\mathcal{N}(H)|\mathbf{a}|_p. \tag{6.16}$$

(b)*If (6.4) is satisfied, then there exist positive constants* $C_1 = C_1(\alpha, a_1, a), c_2 = c_2(\alpha, a_1, a), c_3 = c_3(\alpha, a_1, a)$ *such that for* $n \geqslant C_1 m$,

$$c_2 n^{-q/p'}\|G_n\|_p \leqslant \mathcal{N}(H)|\mathbf{a}|_p \leqslant c_3 n^{-q/p'}\|G_n\|_p. \tag{6.17}$$

The proof requires a number of preparatory results.

Proposition 6.2. *Let H be as in Theorem 6.1(a) and supported on* $\{|\mathbf{x}| \leqslant 1\}$. *Let* $\{\mathbf{y}_j\}_{j=1}^M \subset [-\pi,\pi]^q$, *and m be the smallest integer satisfying the minimal separation condition (6.15). For integer* $n \geqslant 1$ *and* $\mathbf{x} \in [-\pi,\pi]^q$,

$$\sum_{j,|\mathbf{x}-\mathbf{y}_j|\geqslant 1/m} |\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q(m/n)^S\mathcal{N}(H). \tag{6.18}$$

Hence,

$$\sum_{j=1}^M |\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q\left\{1+(m/n)^S\right\}\mathcal{N}(H). \tag{6.19}$$

In particular, if (6.4) is satisfied, then there exists $C_1 = C_1(\alpha, a_1, a) > 0$ *such that for* $n \geqslant C_1 m,\ \ell = 1,\ldots,M$,

$$\sum_{\substack{j=1\\ j\neq\ell}}^M |\Psi_n(H,\mathbf{y}_\ell-\mathbf{y}_j)| \leqslant (1/2)\Psi_n(H,\mathbf{0}). \tag{6.20}$$

Proof. The proof of (6.18) is essentially an integration by parts argument, using an idea used often in the classical theory of polynomial interpolation, given also in [17, Chapter V, Section 9] in a different context. Without loss of generality, we may assume that $\mathcal{N}(H) = 1$. In this proof only, let $\mathbb{Z}_k = \{j : k/m \leqslant |\mathbf{x}-\mathbf{y}_j| \leqslant (k+1)/m\},\ k = 1,2,\ldots,$. We note that since the minimal separation amongst $\mathbf{y}_j$'s does not exceed $1/m$, there are at most ck^{q-1} elements in each $\mathbb{Z}_k$. In view of (6.2), we have

$$\sum_{j,|\mathbf{x}-\mathbf{y}_j|\geqslant 1/m} |\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q \sum_{j,|\mathbf{x}-\mathbf{y}_j|\geqslant 1/m}(n|\mathbf{x}-\mathbf{y}_j|)^{-S} = cn^{q-S}\sum_{k=1}^\infty\sum_{j\in\mathbb{Z}_k}|\mathbf{x}-\mathbf{y}_j|^{-S} \leqslant cn^{q-S}m^S\sum_{k=1}^\infty k^{q-1-S} \leqslant cn^q(m/n)^S,$$

where the convergence of the last series follows from the fact that $S > q$. This proves (6.18).

In light of (6.15), the number of $\mathbf{y}_j$'s with $|\mathbf{x}-\mathbf{y}_j| \leqslant 1/m$ is bounded independently of M and m. Hence, (6.5) implies that

$$\sum_{j,|\mathbf{x}-\mathbf{y}_j|\leqslant 1/m} |\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q.$$

Together with (6.18), this leads to (6.19). The estimate (6.20) follows from (6.5) and (6.18). □

Proposition 6.3. *Let* $S > q$ *be an integer,* $1 \leqslant p \leqslant \infty$, $\{\mathbf{y}_j\}_{j=1}^M \subset [-\pi,\pi]^q$ *and (6.15) be satisfied. For any integer* $n \geqslant 1$, *and* $T \in \mathbb{H}_n^q$, *we have*

$$|(T(\mathbf{y}_1),\ldots,T(\mathbf{y}_M))|_p \leqslant cn^{q/p}\left\{1+(m/n)^S\right\}^{1/p}\|T\|_p. \tag{6.21}$$

Proof. In this proof only, let $h : [0,\infty) \to [0,\infty)$ be a fixed, infinitely differentiable function, $h(t) = 1$ if $0 \leqslant t \leqslant 1/2$, $h(t) = 0$ if $t \geqslant 1$, and we choose $H(\mathbf{x}) := h(|\mathbf{x}|)$. The constants will depend upon this h, but h being fixed in this proof, this dependence need not be specified. A comparison of Fourier coefficients shows that for $T \in \mathbb{H}_{2n}^q$,

$$T(\mathbf{y}) = \frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} T(\mathbf{x})\Psi_{4n}(H,\mathbf{x}-\mathbf{y})d\mathbf{y}.$$

In view of (6.19), we obtain

$$\sum_{j=1}^M |T(\mathbf{y}_j)| \leqslant \|T\|_1 \max_{\mathbf{x}\in[-\pi,\pi]^q}\sum_{j=1}^M |\Psi_{4n}(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q\left\{1+(m/n)^S\right\}\|T\|_1.$$

If $f \in L^1$, we apply this estimate with $\sigma_{2n}(H,f)$ in place of T, and use Corollary 6.1 (with $p = 1$) to deduce that

$$\sum_{j=1}^{M}|\sigma_{2n}(H,f,\mathbf{y}_j)| \leqslant cn^q\left\{1+(m/n)^S\right\}\|f\|_1.$$

In view of Corollary 6.1 (with $p=\infty$), it is clear that for $f\in L^\infty$,

$$\max_{1\leqslant j\leqslant M}|\sigma_{2n}(H,f,\mathbf{y}_j)| \leqslant \|\sigma_{2n}(H,f)\|_\infty \leqslant c\|f\|_\infty.$$

An application of Riesz–Thorin interpolation theorem now implies that for $1\leqslant p<\infty$, and $f\in L^p$,

$$\left\{\sum_{j=1}^{M}|\sigma_{2n}(H,f,\mathbf{y}_j)|^p\right\}^{1/p} \leqslant cn^{q/p}\left\{1+(m/n)^S\right\}^{1/p}\|f\|_p. \tag{6.22}$$

If $T\in\mathbb{H}_n^q$, then a comparison of Fourier coefficients shows that $\sigma_{2n}(H,T)=T$. Therefore, (6.22) implies (6.21). $\square$

Proposition 6.4 below is perhaps well known. A proof can be found in [10, Proposition 6.1].

Proposition 6.4. *Let $M\geqslant 1$ be an integer, $\mathbf{A}$ be an $M\times M$ matrix whose (i,j)th entry is $A_{i,j}$. $1\leqslant p\leqslant\infty$, and $\beta\in[0,1)$. If*

$$\sum_{\substack{i=1\\ i\neq j}}^{M}|A_{j,i}| \leqslant \beta|A_{j,j}|, \quad \sum_{\substack{i=1\\ i\neq j}}^{M}|A_{i,j}| \leqslant \beta|A_{j,j}|, \quad j=1,\ldots,M, \tag{6.23}$$

and $\lambda=\min_{1\leqslant i\leqslant M}|A_{i,i}|>0$, then $\mathbf{A}$ is invertible, and

$$|\mathbf{A}^{-1}\mathbf{b}|_p \leqslant ((1-\beta)\lambda)^{-1}|\mathbf{b}|_p, \quad \mathbf{b}\in\mathbb{R}^M. \tag{6.24}$$

We are now in a position to prove Theorem 6.2.

Proof of Theorem 6.2. Without loss of generality, we may assume that $\mathcal{N}(H)=1$. In view of (6.19), we have for $\mathbf{x}\in[-\pi,\pi]^q$,

$$|G_n(\mathbf{x})| \leqslant \sum_{j=1}^{M}|a_j||\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant |\mathbf{a}|_\infty\sum_{j=1}^{M}|\Psi_n(H,\mathbf{x}-\mathbf{y}_j)| \leqslant cn^q\left\{1+(m/n)^S\right\}|\mathbf{a}|_\infty.$$

Thus,

$$\|G_n\|_\infty \leqslant cn^q\left\{1+(m/n)^S\right\}|\mathbf{a}|_\infty.$$

Using (6.3), we see that

$$\|G_n\|_1 \leqslant \sum_{j=1}^{M}|a_j|\|\Psi_n(H,\circ-\mathbf{y}_j)\|_1 \leqslant c|\mathbf{a}|_1.$$

An application of Riesz–Thorin interpolation theorem with the operator $\mathbf{a}\mapsto\sum_{j=1}^{M}a_j\Psi_n(H,\circ-\mathbf{y}_j)$ implies (6.16).

In the rest of this proof, the constants may depend upon the parameters α, a_1, a in (6.4). In this proof only, let $\mathbf{A}$ be the matrix whose (ℓ,j)th entry is $\Psi_n(H,\mathbf{y}_\ell-\mathbf{y}_j)$ and $\mathbf{b}\in\mathbb{R}^M$ be defined by $b_\ell=G_n(\mathbf{y}_\ell),\ \ell=1,\ldots,M$. In view of (6.20), (6.23) is satisfied with $1/2$ in place of β, and in view of (6.5), we may choose λ to be cn^q. Hence, Proposition 6.4 implies that $\mathbf{A}$ is invertible, and

$$|\mathbf{A}^{-1}\mathbf{b}|_p \leqslant cn^{-q}|\mathbf{b}|_p.$$

Since, $\mathbf{A}^{-1}\mathbf{b}=\mathbf{a}$, we have proved that

$$|\mathbf{a}|_p \leqslant cn^{-q}|\mathbf{b}|_p. \tag{6.25}$$

Since $G_n\in\mathbb{H}_n^q$, we obtain from Proposition 6.3 that

$$|\mathbf{b}|_p = |(G_n(\mathbf{y}_1),\ldots,G_n(\mathbf{y}_M))|_p \leqslant cn^{q/p}\left\{1+(m/n)^S\right\}^{1/p}\|G_n\|_p.$$

Since $n\geqslant C_1m$, this gives

$$|\mathbf{b}|_p \leqslant cn^{q/p}\|G_n\|_p.$$

Together with (6.25), this leads to the second inequality in (6.17). The first inequality follows from (6.16) and the fact that $n\geqslant C_1m$. $\square$

6.2. Sobolev kernel

Our goal in this section is to prove Proposition 4.1 and Theorem 6.3, and establish a few other facts regarding the kernel K_s. In the sequel, we assume $S > q$ is an integer, $h : [0,\infty) \to [0,\infty)$ is a fixed, S times continuously differentiable function, $h(t) = 1$ if $0 \leqslant t \leqslant 1/2, h(t) = 0$ if $t \geqslant 1$, and h is nondecreasing on $[0,\infty)$. We will write $g(t) = h(t) - h(2t)$. Since h is fixed, the dependence of various constants on h need not be indicated. We will simplify our notation and write $\Psi_n(h)$ rather than $\Psi_n(h(|\circ|))$, and similarly for $\Psi_n(g)$. It is elementary calculus to verify that the function $\mathbf{x} \mapsto h(|\mathbf{x}|)$ satisfies all the conditions of Theorem 6.1, including (6.4).

For $s \in \mathbb{R}$, we will write

$$\tilde{\Psi}_{n,s}(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^q} g(|\mathbf{k}|/2^n)(|\mathbf{k}|^2+1)^{-s/2}\exp(i\mathbf{k}\cdot\mathbf{x}). \tag{6.26}$$

The following lemma lists some relevant properties of $\tilde{\Psi}_{n,s}$.

Lemma 6.1. *Let $s \in \mathbb{R}$. We have for integer $n \geqslant 0$,*

$$|\tilde{\Psi}_{n,s}(\mathbf{x})| \leqslant c\frac{2^{n(q-s)}}{\max(1,(2^n|\mathbf{x}|)^S)}, \quad \mathbf{x}\in[-\pi,\pi]^q. \tag{6.27}$$

Further,

$$\max_{\mathbf{x}\in[-\pi,\pi]^q} |\tilde{\Psi}_{n,s}(\mathbf{x})| = \tilde{\Psi}_{n,s}(\mathbf{0}) \sim 2^{n(q-s)}, \tag{6.28}$$

and for $1 \leqslant p \leqslant \infty$,

$$\|\tilde{\Psi}_{n,s}\|_p \leqslant c2^{n(q/p'-s)}. \tag{6.29}$$

Proof. In this proof only, let $g_n(t) = g(t)/(t^2+1/n^2)^{s/2}$. Then for $\mathbf{x}\in[-\pi,\pi]^q$,

$$\tilde{\Psi}_{n,s}(\mathbf{x}) = 2^{-ns}\Psi_{2^n}(g_{2^n}(|\circ|),\mathbf{x}). \tag{6.30}$$

Each $g_n(|\circ|)$ satisfies the conditions of Theorem 6.1. Moreover, using the fact that $g(t) = 0$ for $|t| \leqslant 1/4$, it is easy to verify that $\mathcal{N}(g_n(|\circ|)) \leqslant c$. Therefore, all assertions of the lemma follow directly from Theorem 6.1. $\square$

Proof of Proposition 4.1. Since $s > q/p$, (6.29) used with p' in place of p shows that

$$\sum_{n=0}^{\infty}\|\tilde{\Psi}_{n,s}\|_{p'} \leqslant c\sum_{n=0}^{\infty}2^{n(q/p-s)} < \infty.$$

So, the sequence of trigonometric polynomials, defined by

$$P_N(\mathbf{x}) = 1 + \sum_{n=0}^{N}\tilde{\Psi}_{n,s}(\mathbf{x}) = 1 + \sum_{n=0}^{N}\sum_{\mathbf{k}\in\mathbb{Z}^q} g(|\mathbf{k}|/2^n)(1+|\mathbf{k}|^2)^{-s/2}\exp(i\mathbf{k}\cdot\mathbf{x})$$

converges in $L^{p'}$. All the sums in the above expression being finite sums, we obtain for $N \geqslant 0$,

$$P_N(\mathbf{x}) = 1 + \sum_{\mathbf{k}\in\mathbb{Z}^q}\sum_{n=0}^{N} g(|\mathbf{k}|/2^n)(1+|\mathbf{k}|^2)^{-s/2}\exp(i\mathbf{k}\cdot\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^q} h(|\mathbf{k}|/2^N)(1+|\mathbf{k}|^2)^{-s/2}\exp(i\mathbf{k}\cdot\mathbf{x}).$$

If $\mathbf{k}\in\mathbb{Z}^q$, and $2^N \geqslant 2|\mathbf{k}|$, then $h(|\mathbf{k}|/2^N) = 1$, and $\hat{P}_N(\mathbf{k}) = (1+|\mathbf{k}|^2)^{-s/2}$. Denoting the $L^{p'}$-limiting function of P_N by K_s, it follows that $K_s \in L^{p'}$ and satisfies (4.4). Moreover, $P_N = \sigma_{2^N}(h, K_s)$, and the bound on $\|\tilde{\Psi}_{n,s}\|_{p'}$ in (6.29) used with p' in place of p shows that for $N = 0, 1, 2, \ldots,$

$$\|K_s - \sigma_{2^N}(h,K_s)\|_{p'} \leqslant \sum_{n=N+1}^{\infty}\|\tilde{\Psi}_{n,s}\|_{p'} \leqslant c\sum_{n=N+1}^{\infty}2^{n(q/p-s)} \leqslant c_1 2^{N(q/p-s)}. \tag{6.31}$$

Both sides of the first equation in (4.5) have the same Fourier coefficients, and hence, they are equal almost everywhere. Similarly, a comparison of Fourier coefficients shows that $K_s(-\mathbf{x}) = K_s(\mathbf{x})$ for almost all $\mathbf{x}$. This implies the second equation in (4.5).

For $f \in W_s^p$, a comparison of Fourier coefficients again shows that for integer $m \geqslant 0$,

$$\sigma_{2^m}(h,f,\mathbf{x}) = \frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q}\sigma_{2^m}(h,K_s,\mathbf{x}-\mathbf{y})f^{(s)}(\mathbf{y})d\mathbf{y}.$$

So, (6.31) implies that

$$\|\sigma_1(h,f)\|_\infty \leqslant c\|f^{(s)}\|_p,$$

and

$$\|\sigma_{2^m}(h,f) - \sigma_{2^{m-1}}(h,f)\|_\infty \leqslant \|\sigma_{2^m}(h,K_s) - \sigma_{2^{m-1}}(h,K_s)\|_{p'}\|f^{(s)}\|_p \leqslant c2^{m(q/p-s)}\|f^{(s)}\|_p.$$

Since $s > q/p$, the series $\sigma_1(h,f) + \sum_{m=1}^\infty(\sigma_{2^m}(h,f) - \sigma_{2^{m-1}}(h,f))$ converges uniformly. In this proof only, let F be the limiting function. Then F is continuous, and a calculation shows that $\widehat{F}(\mathbf{k}) = \hat{f}(\mathbf{k})$ for all $\mathbf{k} \in \mathbb{Z}^q$. Therefore, $F = f$ almost everywhere. Choosing the continuous representer in the equivalence class of f to be f, the limit is f. Moreover,

$$\|f\|_\infty \leqslant \|\sigma_1(h,f)\|_\infty + \sum_{m=1}^\infty \|\sigma_{2^m}(h,f) - \sigma_{2^{m-1}}(h,f)\|_\infty \leqslant c\|f^{(s)}\|_p \sum_{m=0}^\infty 2^{m(q/p-s)} \leqslant c_1\|f^{(s)}\|_p.$$

This proves the first estimate in (4.6). The second estimate is proved in the same way. Let L be the smallest integer with $2^L \geqslant n$. Then

$$E_{n,\infty}(f) \leqslant E_{2^L,\infty}(f) \leqslant \sum_{m=L+1}^\infty \|\sigma_{2^m}(h,f) - \sigma_{2^{m-1}}(h,f)\|_\infty \leqslant c\|f^{(s)}\|_p \sum_{m=L+1}^\infty 2^{m(q/p-s)} \leqslant c_1 2^{-L(s-q/p)}\|f^{(s)}\|_p \leqslant c_2 n^{q/p-s}\|f^{(s)}\|_p. \qquad \square$$

Our proof of Theorem 5.1 requires the following theorem that describes an approximation of a typical element of the span of $\{K_s(\circ - \mathbf{y}_j)\}$. We recall that the solution of the minimization problem (1.3) is in this span (with $2s$ in place of s).

Theorem 6.3. *Let $1 \leqslant p \leqslant \infty, s > q/p, \{a_j\}_{j=1}^M \subset \mathbb{R}, G(\mathbf{x}) = \sum_{j=1}^M a_j K_s(\mathbf{x} - \mathbf{y}_j), \mathbf{x} \in [-\pi,\pi]^q$, and $m \geqslant 1$ be the smallest integer such that $\min_{\mathbf{y}_j \neq \mathbf{y}_k}|\mathbf{y}_j - \mathbf{y}_k| \geqslant 1/m$. Then there exists an integer N^*, independent of G, such that $N^* \sim m$ and*

$$\|G - \sigma_{N^*}(h,G)\|_{p'} \leqslant (1/2)\|G\|_{p'}. \tag{6.32}$$

Proof. As in the proof of Lemma 6.1, in this proof only, we write $g_n(t) = g(t)/(t^2 + 1/n^2)^{s/2}$. Then each $g_n(|\circ|)$ satisfies the conditions of Theorem 6.1. Moreover, $\mathcal{N}(g_n(|\circ|)) \leqslant c$, and (6.30) holds. In this proof only, let

$$G_n(\mathbf{x}) := \sum_{j=1}^M a_j \tilde{\Psi}_{n,s}(\mathbf{x} - \mathbf{y}_j) = \sigma_{2^n}(h,G,\mathbf{x}) - \sigma_{2^{n-1}}(h,G,\mathbf{x}), \quad \mathbf{x} \in [-\pi,\pi]^q.$$

Then (6.14) implies that $\|G_n\|_{p'} \leqslant c\|G\|_{p'}$. Moreover, the proof of Proposition 4.1 shows that

$$G(\mathbf{x}) - \sigma_{2^N}(h,G,\mathbf{x}) = \sum_{n=N}^\infty G_n(\mathbf{x}), \tag{6.33}$$

with convergence in the sense of $L^{p'}$.

In view of (6.30), (6.17) applied with $\Psi_{2^n}(g_{2^n})$ and p' in place of p yields that for $n \geqslant \log_2(C_1 m)$,

$$c_2 2^{n(s-q/p)}\|G_n\|_{p'} \leqslant |\mathbf{a}|_{p'} \leqslant c2^{n(s-q/p)}\|G_n\|_{p'} \leqslant c2^{n(s-q/p)}\|G\|_{p'}. \tag{6.34}$$

We now choose L so that 2^L is the smallest power of 2 exceeding $C_1 m$. Then the second inequality in (6.34), used with L in place of n, gives

$$|\mathbf{a}|_{p'} \leqslant cm^{(s-q/p)}\|G_L\|_{p'} \leqslant cm^{(s-q/p)}\|G\|_{p'}. \tag{6.35}$$

From (6.33), (6.34), and (6.35), we conclude that for $2^N \geqslant C_1 m$,

$$\|G - \sigma_{2^N}(h,G)\|_{p'} \leqslant \sum_{n=N}^\infty \|G_n\|_{p'} \leqslant c|\mathbf{a}|_{p'} \sum_{n=N}^\infty 2^{-n(s-q/p)} \leqslant c(m2^{-N})^{(s-q/p)}\|G\|_{p'}.$$

We now choose N so that $2^N \sim m$ and the last term above is at most $(1/2)\|G\|_{p'}$, and set $N^* = 2^N$. $\square$

7. Proofs of the main results in Section 5

We start with the proof of Theorem 5.4.

Proof of Theorem 5.4. In this proof only, we find it convenient to denote the standard inner product on (complex) Euclidean space $\mathbb{C}^n$ by $\langle \circ, \circ \rangle_n$ rather than by the dot notation $\circ \cdot \circ$. The condition (5.5) implies that A^* is injective. Let $W \subseteq \mathbb{C}^D$ be the range of A^*. Since A^* is injective, the formula

$$x^*(A^*\mathbf{c}) = \langle \mathbf{c}, \mathbf{f} \rangle_M.$$

defines a linear functional x^* on W. Using (5.5), we obtain that

$$\sup\{|x^*(A^*\mathbf{c})| : \mathbf{c} \in \mathbb{C}^M, |||A^*\mathbf{c}|||_D^* \leqslant 1\} = \sup\{|\langle \mathbf{c}, \mathbf{f}\rangle_M| : \mathbf{c} \in \mathbb{C}^M, |||A^*\mathbf{c}|||_D^* \leqslant 1\} \leqslant \sup\{|\langle \mathbf{c}, \mathbf{f}\rangle_M| : \mathbf{c} \in \mathbb{C}^M, |||\mathbf{c}|||_M^* \leqslant \kappa\} \leqslant \kappa|||\mathbf{f}|||_M. \tag{7.1}$$

In view of the Hahn–Banach theorem, there exists a norm preserving extension y^* of x^* to $\mathbb{C}^D$ (equipped with the norm $|||\circ|||_D^*$). This extension can be identified with $\mathbf{b} \in \mathbb{C}^D$ by the formula $y^*(\mathbf{d}) = \langle \mathbf{d}, \mathbf{b}\rangle_D$.

Since (7.1) estimates the norm of x^* on W with the norm $|||\circ|||_D^*$ on $\mathbb{C}^D$, the dual norm is $|||\circ|||_D$. The fact that y^* is a norm preserving extension, together with (7.1), shows that $|||\mathbf{b}|||_D \leqslant \kappa|||\mathbf{f}|||_M$.

The fact that y^* is an extension of x^* means that for every $\mathbf{c} \in \mathbb{C}^M$,

$$\langle \mathbf{c}, \mathbf{f}\rangle_M = x^*(A^*\mathbf{c}) = y^*(A^*\mathbf{c}) = \langle A^*\mathbf{c}, \mathbf{b}\rangle_D = \langle \mathbf{c}, A\mathbf{b}\rangle_M.$$

This implies (5.6). $\square$

We are now in a position to prove Theorem 5.1.

Proof of Theorem 5.1. In this proof, we fix n, and denote the nth row of Y by $\{\mathbf{y}_1, \ldots, \mathbf{y}_M\}$. We choose the integer N^* as in Theorem 6.3 used with p' in place of p. Let D be the dimension of $\mathbb{H}^q_{N^*}$, and let A be the $M \times D$ matrix defined by $A_{j,\mathbf{k}} = \exp(i\mathbf{k}\cdot\mathbf{y}_j), |\mathbf{k}| \leqslant N^*,\ j = 1, \ldots, M$. It is not difficult to verify using (6.34) that the expressions

$$|||\mathbf{c}|||_M^* := \left\| \sum_{j=1}^M c_j K_s(\circ - \mathbf{y}_j)\right\|_{p'}, \quad \mathbf{c} \in \mathbb{C}^M, \tag{7.2}$$

and

$$|||\mathbf{d}|||_D := \left\| \sum_{|\mathbf{k}|\leqslant N^*} (|\mathbf{k}|^2 + 1)^{s/2} d_{\mathbf{k}} \exp(i\mathbf{k}\cdot\circ)\right\|_p, \quad \mathbf{d} \in \mathbb{C}^D \tag{7.3}$$

define norms on the Euclidean spaces as indicated. In this proof, if $F \in L^p$, we define

$$F^{(-s)}(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^q} \hat{F}(\mathbf{k})(|\mathbf{k}|^2+1)^{-s/2}\exp(i\mathbf{k}\cdot\mathbf{x}) = \frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} F(\mathbf{y})K_s(\mathbf{x}-\mathbf{y})d\mathbf{y}.$$

We verify the condition (5.5) on the adjoint of A. Let $\mathbf{c} \in \mathbb{C}^M, G(\mathbf{x}) = \sum_{j=1}^M c_j K_s(\mathbf{x} - \mathbf{y}_j)$. Without loss of generality, we may assume that $\|G\|_{p'} \neq 0$. In view of (6.32), there exists $F \in L^p$ such that $\|F\|_p \leqslant 1$ and

$$|||\mathbf{c}|||_M^* = \|G\|_{p'} \leqslant 2\|\sigma_{N^*}(h, G)\|_{p'} \leqslant 2\frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} \sigma_{N^*}(h, G, \mathbf{x})\overline{F(\mathbf{x})}d\mathbf{x}. \tag{7.4}$$

Let $P(\mathbf{x}) = \sigma_{N^*}(h, F^{(-s)}, \mathbf{x}) = \sum_{|\mathbf{k}|\leqslant N^*} d_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{x})$. Then

$$P(\mathbf{y}_j) = (A\mathbf{d})_j, \quad j = 1, \ldots, M. \tag{7.5}$$

Moreover, in view of Corollary 6.1 and the definition (7.3), we deduce that

$$|||\mathbf{d}|||_D = \|\sigma_{N^*}(F)\|_p \leqslant c\|F\|_p \leqslant c. \tag{7.6}$$

Finally, a straightforward calculation shows (keeping in mind that all the sums below are finite sums) that

$$\begin{aligned}\frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} \sigma_{N^*}(h, G, \mathbf{x})\overline{F(\mathbf{x})}d\mathbf{x} &= \sum_{\mathbf{k}\in\mathbb{Z}^q} h\left(\frac{\mathbf{k}}{N^*}\right)\sum_{j=1}^M c_j(|\mathbf{k}|^2+1)^{-s}\exp(-i\mathbf{k}\cdot\mathbf{y}_j)\overline{\frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} F(\mathbf{x})\exp(-i\mathbf{k}\cdot\mathbf{x})d\mathbf{x}} \\ &= \sum_{j=1}^M c_j \sum_{\mathbf{k}\in\mathbb{Z}^q} h\left(\frac{\mathbf{k}}{N^*}\right)(|\mathbf{k}|^2+1)^{-s}\overline{\widehat{F}(\mathbf{k})}\exp(-i\mathbf{k}\cdot\mathbf{y}_j) \\ &= \sum_{j=1}^M c_j \sum_{\mathbf{k}\in\mathbb{Z}^q} \overline{\sigma_{N^*}(\widehat{h, F^{(-s)}})(\mathbf{k})}\exp(-i\mathbf{k}\cdot\mathbf{y}_j) = \sum_{j=1}^M c_j\overline{P(\mathbf{y}_j)}.\end{aligned} \tag{7.7}$$

Together with (7.4), the last three equations show that for any $\mathbf{c} \in \mathbb{C}^M$,

$$\frac{|||\mathbf{c}|||_M^*}{2} \leqslant \frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q} \sigma_{N^*}(h, G, \mathbf{x})\overline{F(\mathbf{x})}d\mathbf{x} = \sum_{j=1}^M c_j\overline{P(\mathbf{y}_j)} = \sum_{j=1}^M c_j\overline{(A\mathbf{d})_j} = \sum_{|\mathbf{k}|\leqslant N^*} (A^*\mathbf{c})_{\mathbf{k}}\overline{d_{\mathbf{k}}} \leqslant |||A^*\mathbf{c}|||_D^*|||\mathbf{d}|||_D \leqslant c|||A^*\mathbf{c}|||_D^*.$$

Thus, the matrix A satisfies the conditions of Theorem 5.4 with $\kappa := 2c$.

We now choose $T \in \mathbb{H}^q_{N^*}$ such that

$$\|f - T\|_{W_s^p} \leqslant 2\inf\{\|f - T\|_{W_s^p} : T \in \mathbb{H}^q_{N^*}\}.$$

Let $\mathbf{f}\in\mathbb{C}^M$ be the vector defined by $f_j=f(\mathbf{y}_j)-T(\mathbf{y}_j),\ j=1,\ldots,M$. In view of Theorem 5.4, there exists $\mathbf{b}\in\mathbb{C}^D$ such that $A\mathbf{b}=\mathbf{f}$, and

$$|||\mathbf{b}|||_D\leqslant\kappa|||\mathbf{f}|||_M. \tag{7.8}$$

Let $Q\in\mathbb{H}_N^*$ be defined by $Q(\mathbf{x})=\sum_{\mathbf{k}}b_{\mathbf{k}}\exp(i\mathbf{k}\cdot\mathbf{x})$. The definition (7.3) is designed so that $|||\mathbf{b}|||_D=\|Q^{(s)}\|_p=\|Q\|_{W_s^p}$. We need to estimate $|||\mathbf{f}|||_M$. Using the definition of the dual norm (7.2), we find $\mathbf{c}\in\mathbb{C}^M$ such that

$$|||\mathbf{f}|||_M\leqslant 2\sum_{j=1}^M c_j\overline{(f(\mathbf{y}_j)-T(\mathbf{y}_j))},$$

and

$$|||\mathbf{c}|||_M^*=\left\|\sum_{j=1}^M c_jK_s(\circ-\mathbf{y}_j)\right\|_{p'}\leqslant 1.$$

In view of the representation identity (4.5), and recalling that K_s is a real valued and symmetric kernel, we have

$$\frac{1}{2}|||\mathbf{f}|||_M\leqslant\sum_{j=1}^M c_j\overline{(f(\mathbf{y}_j)-T(\mathbf{y}_j))}=\frac{1}{(2\pi)^q}\int_{[-\pi,\pi]^q}\left\{\sum_{j=1}^M c_jK_s(\mathbf{x}-\mathbf{y}_j)\right\}\{f^{(s)}(\mathbf{x})-T^{(s)}(\mathbf{x})\}d\mathbf{x}\leqslant\left\|\sum_{j=1}^M c_jK_s(\circ-\mathbf{y}_j)\right\|_{p'}\|f-T\|_{W_s^p}$$
$$\leqslant\|f-T\|_{W_s^p}.$$

Therefore, (7.8) implies that

$$\|Q\|_{W_s^p}\leqslant 2\kappa\|f-T\|_{W_s^p}. \tag{7.9}$$

We now define $\mathbf{P}(f,\mathbf{x})=T(\mathbf{x})+Q(\mathbf{x})$. It is easy to verify that

$$\mathbf{P}(f,\mathbf{y}_j)=T(\mathbf{y}_j)+Q(\mathbf{y}_j)=T(\mathbf{y}_j)+f_j=T(\mathbf{y}_j)+f(\mathbf{y}_j)-T(\mathbf{y}_j)=f(\mathbf{y}_j),\quad j=1,\ldots,M,$$

In view of our choice of T and (7.9)

$$\|f-\mathbf{P}(f)\|_{W_s^p}\leqslant\|f-T\|_{W_s^p}+2\kappa\|f-T\|_{W_s^p}\leqslant(2+4\kappa)\inf\{\|f-T\|_{W_s^p}:\ T\in\mathbb{H}_{N^*}^q\}.\qquad\square$$

In order to deduce Theorem 5.2 from Theorem 5.1, we need the following lemma.

Lemma 7.1. *For integer $n\geqslant 1, 1\leqslant p\leqslant\infty$, and $T\in\mathbb{H}_n^q$, we have*

$$\left\{\frac{1}{n^q}\sum_{0\leqslant\mathbf{k}\leqslant 3n-1}|T(2\pi\mathbf{k}/(3n))|^p\right\}^{1/p}\sim\|T\|_p. \tag{7.10}$$

Proof. When $q=1$, (7.10) is the classical Marcinkiewicz–Zygmund inequality [17, Chapter X, Theorems 7.5 and 7.28]. If $T\in\mathbb{H}_n^q$, then T is a trigonometric polynomial of coordinatewise order at most n. So, in the case when $q>1$, one obtains (7.10) by applying its univariate version to each of the variables separately. $\square$

Proof of Theorem 5.2. In view of Theorem 5.1 and Lemma 7.1, we have

$$\left\{\frac{1}{N^{*q}}\sum_{0\leqslant\mathbf{k}\leqslant 3N^*-1}|\mathbf{P}(f)^{(s)}(2\pi\mathbf{k}/(3N^*))|^p\right\}^{1/p}\leqslant c\|\mathbf{P}(f)\|_{W_s^p}\leqslant c\|f\|_{W_s^p}.$$

So, the minimization problem (5.3) has a feasible solution. Since $\mathbb{H}_{N^*}^q$ is a finite dimensional space, this implies that the problem has a solution. $\square$

The proof of Theorem 5.3 uses the following lemma. This lemma is proved in much greater generality in [8, Theorem 3.2, Chapter 15].

Lemma 7.2. *Let $1\leqslant p\leqslant\infty$, $s'>0$. Then for any $c>0$, the set*

$$B_{c,s',p}:=\left\{f\in L^p:\sup_{n\geqslant 1}2^{ns'}E_{n,p}(f)\leqslant c\right\}$$

is compact in L^p.

Proof of Theorem 5.3. We observe that in view of (4.6) and the fact that $\|\mathbb{T}_n^*\|_{W_s^p} \leqslant c\|f\|_{W_s^p}$ for all n, the sequence $\{\mathbb{T}_n^*\} \subset B_{c,s-q/p,\infty}$ for a suitable constant c. Let Λ_1 be any subsequence of Λ. Then Lemma 7.2 shows that the sequence $\{\mathbb{T}_n^*\}_{n\in\Lambda_1}$ has a subsequence $\{\mathbb{T}_n^*\}_{n\in\Lambda_2}$, which converges uniformly. Let P be the limit of this subsequence. We will show that if (5.4) is satisfied, then $P(\mathbf{x}_0) = f(\mathbf{x}_0)$. Let $\epsilon > 0$ be arbitrary. Since P and f are continuous on $[-\pi,\pi]^q$, there is $\delta > 0$ such that

$$|f(\mathbf{x}) - f(\mathbf{y})| \leqslant \epsilon/3, \quad |P(\mathbf{x}) - P(\mathbf{y})| \leqslant \epsilon/3, \quad \text{for all } \mathbf{x},\mathbf{y} \in [-\pi,\pi]^q, \quad |\mathbf{x}-\mathbf{y}| \leqslant \delta.$$

Further, there exists N so that $n \geqslant N, n \in \Lambda_2$ imply that $\|P - \mathbb{T}_n^*\|_\infty \leqslant \epsilon/3$. In view of (5.4), there exists $n \in \Lambda_2, n \geqslant N$ such that some point $\mathbf{y}_{j,n} \in Y$ satisfies $|\mathbf{y}_{j,n} - \mathbf{x}_0| \leqslant \delta$. Then $f(\mathbf{y}_{j,n}) = \mathbb{T}_n^*(\mathbf{y}_{j,n})$, and we have

$$\begin{aligned}|f(\mathbf{x}_0) - P(\mathbf{x}_0)| &\leqslant |f(\mathbf{x}) - f(\mathbf{y}_{j,n})| + |f(\mathbf{y}_{j,n}) - \mathbb{T}_n^*(\mathbf{y}_{j,n})| + |\mathbb{T}_n^*(\mathbf{y}_{j,n}) - P(\mathbf{y}_{j,n})| + |P(\mathbf{y}_{j,n}) - P(\mathbf{x}_0)| \leqslant \epsilon/3 + 0 + \|\mathbb{T}_n^* - P\|_\infty + \epsilon/3 \\ &\leqslant \epsilon.\end{aligned}$$

Since this is true for every subsequential limit of $\mathbb{T}_n^*, n \in \Lambda$, this completes the proof of the theorem. □

References

[1] B. Adcock, A. Hansen, A generalized sampling theorem for stable reconstructions in arbitrary bases, J. Fourier Anal. Appl. (2012) 1–32.
[2] N. Castro-Gonzalez, J. Ceballos, F.M. Dopico, J.M. Molera, Multiplicative perturbation theory and accurate solution of least squares problems, submitted for publication.
[3] S. Chandrasekaran, K.R. Jayaraman, J. Moffitt, H.N. Mhaskar, S. Pauli, Minimum Sobolev norm schemes and applications in image processing, in: Proceedings of the IS&T/SPIE Symposium on Electronic Imaging: Science & Technology, San José, CA, January 2010.
[4] S. Chandrasekaran, K.R. Jayaraman, M. Gu, H.N. Mhaskar, J. Moffitt, Higher order numerical discretization methods with Sobolev norm minimization, in: Proceedings of the 11th International Conference on Computational Sciences, Singapore, June 2011.
[5] J. Czipszer, G. Freud, Sur l'approximation d'uns fonction periodique et ses derivees successives par un polynomie trigonometrique et par ses derivees successives, Acta Math. 5 (1957) 285–290.
[6] M. Golomb, H.F. Weinberger, Optimal approximation and error bounds, in: R.E. Langer, (Ed.), Proceedings of a Symposium on numerical approximation, Publication No. 1 of the Mathematics Research Center, U.S. Army, The University of Wisconsin, The University of Wisconsin Press, Madison, Wis, April 21–23, 1958, pp. 117–190.
[7] P.D. Hough, S.A. Vavasis, Complete orthogonal decomposition for weighted least squares, SIAM J. Matrix Anal. Appl. 18 (2) (April 1997) 369–392.
[8] G.G. Lorentz, M.V. Golitschek, Y. Makovoz, Constructive Approximation, Advanced Problems, Springer-Verlag, New York, 1996.
[9] H.N. Mhaskar, Weighted quadrature formulas and approximation by zonal function networks on the sphere, J. Complexity 22 (2006) 348–370.
[10] H.N. Mhaskar, Eignets for function approximation on manifolds, Appl. Comput. Harm. Anal. 29 63–87, in press.
[11] H.N. Mhaskar, F. Narcowich, J. Prestin, J.D. Ward, L^p Bernstein estimates and approximation by spherical basis functions, Math. Comput. 79 (271) (2010) 1647–1679.
[12] H.N. Mhaskar, F. Narcowich, N. Sivakumar, J.D. Ward, Approximation with interpolatory constraints, Proc. Amer. Math. Soc. 130 (5) (2002) 1355–1364.
[13] H.N. Mhaskar, J. Prestin, On the detection of singularities of a periodic function, Adv. Comput. Math. 12 (2000) 95–131.
[14] I.P. Natanson, Constructive Function Theory, 3, Frederick Ungar Publications, New York, 1965.
[15] E.M. Stein, G. Weiss, Introduction to Fourier Analysis on Euclidean Spaces, Princeton University Press, Princeton, 1990.
[16] J. Szabados, P. Vértesi, Interpolation of Functions, World Scientific Publishing Co., Singapore, 1990.
[17] A. Zygmund, Trigonometric Series, Cambridge University Press, Cambridge, 1977.