

Stability Results for Scattered Data Interpolation by Trigonometric Polynomials

Stefan Kunis and Daniel Potts *

A fast and reliable algorithm for the optimal interpolation of scattered data on the torus \mathbb{T}^d by multivariate trigonometric polynomials is presented. The algorithm is based on a variant of the conjugate gradient method in combination with the fast Fourier transforms for nonequispaced nodes. The main result is that under mild assumptions the total complexity for solving the interpolation problem at M arbitrary nodes is of order $\mathcal{O}(M \log M)$. This result is obtained by the use of localised trigonometric kernels where the localisation is chosen in accordance to the spatial dimension d . Numerical examples show the efficiency of the new algorithm.

2000 *Mathematics Subject Classification.* 65F10, 65T40, 65F15.

Key words and phrases. trigonometric approximation, scattered data interpolation, localisation, iterative methods, fast Fourier transforms for nonequispaced nodes, FFT

1 Introduction

We discuss the approximation of scattered data by d -variate periodic functions $f : \mathbb{T}^d \rightarrow \mathbb{C}$, where $\mathbb{T} := [-\frac{1}{2}, \frac{1}{2})$ denotes the torus. In practical applications we are often confronted with the situation that experimental data or measured values of a function are only known at a finite sampling set $\mathcal{X} := \{\mathbf{x}_j \in \mathbb{T}^d : j = 0, \dots, M-1\}$. Especially, nonuniform sampling sets appear in more and more applications in recent years. Given a notion of the distance of two points by $\text{dist}(\mathbf{x}, \mathbf{x}_0) := \min_{\mathbf{j} \in \mathbb{Z}^d} \|(\mathbf{x} + \mathbf{j}) - \mathbf{x}_0\|_\infty$, we measure the “nonuniformity” of \mathcal{X} by the mesh norm and the separation distance, defined by

$$\delta := 2 \max_{\mathbf{x} \in \mathbb{T}^d} \min_{j=0, \dots, M-1} \text{dist}(\mathbf{x}_j, \mathbf{x}), \quad q := \min_{j, l=0, \dots, M-1; j \neq l} \text{dist}(\mathbf{x}_j, \mathbf{x}_l),$$

respectively. Obviously, the relation $q \leq M^{-1/d} \leq \delta$ is fulfilled.

*Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
({kunis,potts}@mathematik.tu-chemnitz.de)

For given samples $(\mathbf{x}_j, y_j) \in \mathbb{T}^d \times \mathbb{C}$, $j = 0, \dots, M-1$, a polynomial degree $N \in 2\mathbb{N}$, and the index set $I_N := \{-\frac{N}{2}, \dots, \frac{N}{2} - 1\}^d$ of frequencies, we construct a d -variate trigonometric polynomial

$$f(\mathbf{x}) := \sum_{\mathbf{k} \in I_N} \hat{f}_{\mathbf{k}} e^{2\pi i \mathbf{k} \mathbf{x}}$$

such that $f(\mathbf{x}_j) \approx y_j$, $j = 0, \dots, M-1$. Turning this into matrix vector notation, we aim to solve the system of linear equations

$$\mathbf{A} \hat{\mathbf{f}} \approx \mathbf{y} \tag{1.1}$$

for the unknown vector of Fourier coefficients $\hat{\mathbf{f}} := (\hat{f}_{\mathbf{k}})_{\mathbf{k} \in I_N} \in \mathbb{C}^{N^d}$. Throughout the paper, we denote the vector of the given sample values by $\mathbf{y} := (y_j)_{j=0, \dots, M-1} \in \mathbb{C}^M$ and the nonequispaced Fourier matrix by

$$\mathbf{A} = \mathbf{A}_{\mathcal{X}} := \left(e^{2\pi i \mathbf{k} \mathbf{x}_j} \right)_{j=0, \dots, M-1; \mathbf{k} \in I_N} \in \mathbb{C}^{M \times N^d}.$$

In contrast to the widely used nonequispaced FFT for the fast matrix vector multiplication with \mathbf{A} , see [16] and its references, the *efficient* solution of (1.1) is still a challenging goal. Besides recently developed sparse reconstruction techniques, see e.g. [5, 13] and their references, a standard method to determine $\hat{\mathbf{f}}$ is to solve the general linear least squares problem $\|\hat{\mathbf{f}}\|_2 \rightarrow \min$ subject to $\|\mathbf{y} - \mathbf{A} \hat{\mathbf{f}}\|_2 = \min$, see, e.g., [3, p. 15]. This can be done by means of the singular value decomposition which is not practical in the present situation for large problems due to its time and memory requirements. Direct solvers for the univariate case $d = 1$ in [6, 18] obtain a solution in $\mathcal{O}(NM)$ floating point operations.

For $N^d < M$, the linear system (1.1) is over-determined, so that in general the given data \mathbf{y} will be only approximated up to a *residual* $\mathbf{r} := \mathbf{y} - \mathbf{A} \hat{\mathbf{f}}$. In order to compensate for clusters in the sampling set \mathcal{X} , it is useful to incorporate weights $w_j > 0$ and to consider the weighted approximation problem

$$\|\mathbf{y} - \mathbf{A} \hat{\mathbf{f}}\|_{\mathbf{W}}^2 = \sum_{j=0}^{M-1} w_j |y_j - f(\mathbf{x}_j)|^2 \xrightarrow{\hat{\mathbf{f}}} \min, \tag{1.2}$$

where $\mathbf{W} := \text{diag}(w_j)_{j=0, \dots, M-1}$. In [10] it has been proven that this problem has a unique solution if $N < (\frac{\pi}{\log 2} d \delta)^{-1}$. Its solution is computed iteratively by means of the conjugate gradient method in [7, 2, 9], where the multilevel Toeplitz structure of $\mathbf{A}^H \mathbf{W} \mathbf{A}$ is used for fast matrix vector multiplications. Slightly more stable with respect to rounding errors is the CGNR method, cf. [3, pp. 288], which iterates the original residual $\mathbf{r}_l = \mathbf{y} - \mathbf{A} \hat{\mathbf{f}}_l$ instead of the residual $\mathbf{A}^H \mathbf{W} \mathbf{r}_l$ of the normal equations. Note furthermore, that it has been suggested in [17] to incorporate some “knowledge on the decay of the Fourier coefficients” $\hat{\mathbf{W}} := \text{diag}(\hat{w}_{\mathbf{k}})_{\mathbf{k} \in I_N}$, $\hat{w}_{\mathbf{k}} > 0$. Their approach is based on the weighted least squares problem

$$\left\| \mathbf{A}^H \hat{\mathbf{W}} (\mathbf{y} - \mathbf{A} \hat{\mathbf{f}}) \right\|_{\hat{\mathbf{W}}^{-1}} \xrightarrow{\hat{\mathbf{f}}} \min.$$

In contrast, we focus on the under-determined and consistent linear system $\mathbf{A}\hat{\mathbf{f}} = \mathbf{y}$, i.e., we expect to interpolate the given data $y_j \in \mathbb{C}$, $j = 0, \dots, M - 1$, exactly. We show that the nonequispaced Fourier matrix \mathbf{A} has full rank M for every polynomial degree $N > 2d q^{-1}$. In particular, we incorporate *damping factors* $\hat{w}_{\mathbf{k}} > 0$, $\mathbf{k} \in I_N$, and consider the *optimal interpolation problem*

$$\|\hat{\mathbf{f}}\|_{\hat{\mathbf{W}}^{-1}}^2 = \sum_{\mathbf{k} \in I_N} \frac{|\hat{f}_{\mathbf{k}}|^2}{\hat{w}_{\mathbf{k}}} \xrightarrow{\hat{\mathbf{f}}} \min \quad \text{subject to} \quad \mathbf{A}\hat{\mathbf{f}} = \mathbf{y}, \quad (1.3)$$

where $\hat{\mathbf{W}} := \text{diag}(\hat{w}_{\mathbf{k}})_{\mathbf{k} \in I_N}$. We prove that for a large class of “smooth” damping factors $\hat{w}_{\mathbf{k}}$ problem (1.3) is well-conditioned, where the “smoothness” has to be chosen with respect to the spatial dimension d . We propose to solve problem (1.3) by a version of the conjugate gradient method in combination with the nonequispaced FFT [19, 16, 12] to efficiently perform each iteration step.

The outline of this paper is as follows: In Section 2 we set up the basic notation and relate the optimal interpolation problem (1.3) to a particular trigonometric kernel. Furthermore, we propose Algorithm 1 for computing the solution of the interpolation problem efficiently. For the sake of analysing the convergence of this algorithm, we then present our theory on localised trigonometric kernels in Section 3. Our first result in Theorem 3.3 is a version of the typical smoothness-decay principle in Fourier analysis and relates the “smoothness” of the weights in (1.3) to the localisation of the corresponding trigonometric kernel. We use this decay in Section 4 to prove that well separated sampling nodes yield a stable interpolation problem (1.3). The eigenvalue estimates are given for the univariate setting in Theorem 4.1 and for the multivariate setting in Theorem 4.6. Subsequently, Corollary 4.7 applies the general result to a particular class of damping factors and concludes with conditions sufficient for the full rank of \mathbf{A} . As the equidistant case in Theorem 4.10 and Corollary 4.11 reveals, the assumption on the separation distance is of optimal order. We provide numerical examples in Section 5 and draw our conclusion in Section 6. The software and all numerical examples are available from our NFFT-homepage [12].

2 Optimal interpolation and its iterative solution

After setting up our notation in Definition 2.1, we prove in Lemma 2.2 that the optimal interpolation problem (1.3) can be stated as normal equations and the matrix in these equations obeys special structure. Furthermore, we propose Algorithm 1 for the iterative solution of the interpolation problem and state a basic convergence result for this scheme.

Definition 2.1. Let $d \in \mathbb{N}$, $N \in 2\mathbb{N}$, and $I_N = \{-\frac{N}{2}, \dots, \frac{N}{2} - 1\}^d$ be given. We define for positive weights $\hat{w}_{\mathbf{k}} > 0$, $\mathbf{k} \in I_N$, with normalisation $\sum_{\mathbf{k} \in I_N} \hat{w}_{\mathbf{k}} = 1$ and for $\mathbf{x} \in \mathbb{T}^d$ the trigonometric *kernel*

$$K_N(\mathbf{x}) := \sum_{\mathbf{k} \in I_N} \hat{w}_{\mathbf{k}} e^{2\pi i \mathbf{k} \mathbf{x}}.$$

The particular class of *tensor product kernels* is given by

$$K_N(\mathbf{x}) = \prod_{t=0}^{d-1} \tilde{K}_N(x_t)$$

where \tilde{K}_N denotes a univariate kernel and $\mathbf{x} = (x_0, \dots, x_{d-1})^\top$.

Furthermore, given a sampling set $\mathcal{X} \subset \mathbb{T}^d$, we define the *kernel matrix*

$$\mathbf{K}_N := (K_N(\mathbf{x}_j - \mathbf{x}_l))_{j,l=0,\dots,M-1} \in \mathbb{C}^{M \times M}. \quad (2.1)$$

We denote by $\Lambda = \Lambda(\mathbf{K}_N)$ and $\lambda = \lambda(\mathbf{K}_N)$ the largest and smallest eigenvalue of the kernel matrix \mathbf{K}_N , respectively. Their ratio is denoted by the condition number $\text{cond}(\mathbf{K}_N) = \frac{\Lambda}{\lambda}$. \square

Note, that from the definition immediately follows $K_N(\mathbf{0}) = \max_{\mathbf{x} \in \mathbb{T}^d} |K_N(\mathbf{x})| = 1$ and $(\mathbf{K}_N)_{j,j} = 1$, $j = 0, \dots, M-1$. The following theorem collects some basic facts.

Lemma 2.2. *Let the number of nodes $M \in \mathbb{N}$, the sampling set $\mathcal{X} \subset \mathbb{T}^d$, the polynomial degree $N \in 2\mathbb{N}$, and the damping factors $\hat{w}_{\mathbf{k}} > 0$, $\mathbf{k} \in I_N$, be given. The optimal interpolation problem (1.3) is equivalent to the damped normal equations of second kind*

$$\mathbf{K}_N \tilde{\mathbf{f}} = \mathbf{y}, \quad \hat{\mathbf{f}} = \hat{\mathbf{W}} \mathbf{A}^H \tilde{\mathbf{f}}, \quad (2.2)$$

where the kernel matrix $\mathbf{K}_N \in \mathbb{C}^{M \times M}$ obeys the factorisation

$$\mathbf{K}_N = \mathbf{A} \hat{\mathbf{W}} \mathbf{A}^H, \quad (2.3)$$

hence is positive semidefinite.

Proof. The second assertion follows from $(\mathbf{A} \hat{\mathbf{W}} \mathbf{A}^H)_{j,l} = \sum_{\mathbf{k} \in I_N} e^{2\pi i \mathbf{k} \mathbf{x}_j} \hat{w}_{\mathbf{k}} e^{-2\pi i \mathbf{k} \mathbf{x}_l}$ and Definition 2.1. Furthermore, a solution $\hat{\mathbf{f}}$ of $\mathbf{A} \hat{\mathbf{f}} = \mathbf{y}$ has minimal weighted norm if and only if it is perpendicular with respect to the weights to the null-space of \mathbf{A} , i.e., $\hat{\mathbf{W}}^{-1/2} \hat{\mathbf{f}} \perp \mathcal{N}(\mathbf{A} \hat{\mathbf{W}}^{1/2})$. We conclude (2.2) by the fact that the orthogonal complement of the null-space of a matrix is just the range of its adjoint. \blacksquare

Denoted in Algorithm 1 by CGNE, cf. [3, pp. 288], we solve the **N**ormal equations (2.2) by the **C**onjugate **G**radient method, minimising in each iteration the **E**rror.

The proposed method finds approximations from a Krylov space closely related to the one of the CGNR method for (1.2), but with minimal error instead of minimal residual. Note that we exploit the factorisation in (2.3) to iterate the original vector $\hat{\mathbf{f}}$ instead of the vector $\tilde{\mathbf{f}}$, cf. equation (2.2). Hence, we use fast matrix vector multiplications for \mathbf{A} and \mathbf{A}^H by means of the fast Fourier transforms at nonequispaced nodes (NFFT) having an arithmetical complexity of $\mathcal{O}(N^d \log(N^d) + M |\log \epsilon|^d)$ in each iteration, where ϵ is the prescribed accuracy. Details concerning NFFT algorithms can be found for example in [19, 16] and a corresponding software package in [12]. Applying the standard estimate for the convergence of the conjugate gradient method we obtain the following lemma.

Algorithm 1 CGNE

Input: dimension $d \in \mathbb{N}$, number of samples $M \in \mathbb{N}$, polynomial degree $N \in 2\mathbb{N}$;
sampling set $\mathcal{X} \subset \mathbb{T}^d$, samples $\mathbf{y} \in \mathbb{C}^M$, and initial vector $\hat{\mathbf{f}}_0 \in \mathbb{C}^{N^d}$

$$\mathbf{r}_0 = \mathbf{y} - \mathbf{A}\hat{\mathbf{f}}_0$$

$$\hat{\mathbf{p}}_0 = \mathbf{A}^H \mathbf{r}_0$$

for $l = 0, \dots$ do

$$\alpha_l = \mathbf{r}_l^H \mathbf{r}_l / \hat{\mathbf{p}}_l^H \hat{\mathbf{W}} \hat{\mathbf{p}}_l$$

$$\hat{\mathbf{f}}_{l+1} = \hat{\mathbf{f}}_l + \alpha_l \hat{\mathbf{W}} \hat{\mathbf{p}}_l$$

$$\mathbf{r}_{l+1} = \mathbf{r}_l - \alpha_l \mathbf{A} \hat{\mathbf{W}} \hat{\mathbf{p}}_l$$

$$\beta_l = \mathbf{r}_{l+1}^H \mathbf{r}_{l+1} / \mathbf{r}_l^H \mathbf{r}_l$$

$$\hat{\mathbf{p}}_{l+1} = \beta_l \hat{\mathbf{p}}_l + \mathbf{A}^H \mathbf{r}_{l+1}$$

end for

Output: the l -th iterate $\hat{\mathbf{f}}_l$

Lemma 2.3. *Let the kernel matrix \mathbf{K}_N in (2.1) be regular and let $\hat{\mathbf{e}}_l := \hat{\mathbf{f}}_l - \hat{\mathbf{W}} \mathbf{A}^H \mathbf{K}_N^{-1} \mathbf{y}$ denote the error of the l -th iterate within Algorithm 1. Then the a-priori error bound*

$$\|\hat{\mathbf{e}}_l\|_{\hat{\mathbf{W}}^{-1}} \leq 2 \left(\frac{\sqrt{\Lambda} - \sqrt{\lambda}}{\sqrt{\Lambda} + \sqrt{\lambda}} \right)^l \|\hat{\mathbf{e}}_0\|_{\hat{\mathbf{W}}^{-1}}$$

holds true.

Proof. We note that $\|\hat{\mathbf{e}}_l\|_{\hat{\mathbf{W}}^{-1}} = \|\tilde{\mathbf{f}}_l - \mathbf{K}_N^{-1} \mathbf{y}\|_{\mathbf{K}_N}$, where $\tilde{\mathbf{f}}_l$ denotes the l -th iterate of the conjugate gradient method applied to equation $\mathbf{K}_N \tilde{\mathbf{f}} = \mathbf{y}$, cf. Lemma 2.2, and apply the standard estimate for the conjugate gradient method, see also [3, pp. 288]. ■

This result includes the special case of $M = N^d$ equispaced nodes and no damping, i.e. $\hat{w}_{\mathbf{k}} = 1$, $\mathbf{k} \in I_N$, where the first iterate of our algorithm is already the solution to equation (2.2). We present estimates for the extremal eigenvalues λ, Λ dependent only on the quantities N, q , and the damping factors $\hat{w}_{\mathbf{k}}$, $\mathbf{k} \in I_N$. Analogous results for the stability of the interpolation by radial and zonal functions are obtained in [15, 20]. Section 3 prepares our estimates by constructing localised kernels.

Remark 2.4. Before that, we would like to comment on the following:

The weighted norm in (1.3) is induced by the inner product $\hat{\mathbf{g}}^H \hat{\mathbf{W}}^{-1} \hat{\mathbf{f}}$. In particular, the definition $\langle f, g \rangle_{\hat{\mathbf{W}}^{-1}} := \hat{\mathbf{g}}^H \hat{\mathbf{W}}^{-1} \hat{\mathbf{f}}$ makes the space of trigonometric polynomials

$T_N := \text{span} \{e^{2\pi i \mathbf{k} \cdot} : \mathbf{k} \in I_N\}$ to a reproducing kernel Hilbert space. Its reproducing kernel is given by K_N , i.e., the point evaluations obey $f(\mathbf{x}) = \langle f, K_N(\cdot - \mathbf{x}) \rangle_{\hat{\mathbf{W}}^{-1}}$.

Moreover, the solution $f(\mathbf{x}) = \sum_{\mathbf{k} \in I_N} \hat{f}_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}}$ of the normal equations (2.2) has comparable norm to the given samples, i.e.,

$$\Lambda^{-1} \|\mathbf{y}\|_2^2 \leq \langle f, g \rangle_{\hat{\mathbf{W}}^{-1}} \leq \lambda^{-1} \|\mathbf{y}\|_2^2.$$

This norm equivalence is due to fact, that the field of values of the matrix \mathbf{K}_N^{-1} is bounded by its extremal eigenvalues and $\mathbf{y}^H \mathbf{K}_N^{-1} \mathbf{y} = \tilde{\mathbf{f}}^H \mathbf{K}_N \tilde{\mathbf{f}} = \hat{\mathbf{f}}^H \hat{\mathbf{W}}^{-1} \hat{\mathbf{f}}$. \square

3 Localised kernels

Starting from a class of admissible weight functions in Definition 3.1, we construct localised trigonometric kernels in Theorem 3.3, where Lemma 3.2 serves as an intermediate step. Following the smoothness-decay principle in Fourier analysis, we relate the smoothness of the weight function to the decay of the kernel K_N built upon the sampled weights. A related approach is taken in [14, Thm. 2.2] for the detection of singularities. The particular class of B-Spline kernels, cf. Definition 3.4, is considered in Corollary 3.5. While we present our results on the connection between smooth weight functions and localised kernels for the univariate case, we give its generalisation to the class of tensor product kernels in Corollary 3.7.

Definition 3.1. For $\beta \in \mathbb{N}$, $\beta \geq 2$, a continuous function $g : \mathbb{R} \rightarrow \mathbb{R}$ is an *admissible weight function of order β* if it is nonnegative, possesses a $(\beta - 1)$ -fold derivative $g^{(\beta-1)}$ of bounded variation, i.e.,

$$\left| g^{(\beta-1)} \right|_V := \int_{\mathbb{R}} \left| dg^{(\beta-1)}(z) \right| = \sup \sum_{j=0}^{n-1} \left| g^{(\beta-1)}(z_{j+1}) - g^{(\beta-1)}(z_j) \right| < \infty,$$

where the supremum is taken over all strictly increasing real sequences $\{z_j\}_{j \in \mathbb{N}_0}$, and satisfies the additional properties $\text{supp } g = [-\frac{1}{2}, \frac{1}{2}]$, $g^{(\gamma)}(\pm \frac{1}{2}) = 0$ for $\gamma = 0, \dots, \beta - 1$, $g(z) > 0$, $|z| < \frac{1}{2}$, and the normalisation $\|g\|_{L^1} = 1$. We denote by $BV_0^{\beta-1}$ the set of admissible weight functions of order β .

Furthermore, we define for notational convenience the *zeta function* $\zeta(\beta) := \sum_{r=1}^{\infty} r^{-\beta}$, $\beta > 1$, and for $g \in BV_0^{\beta-1}$ the *norm of the samples*

$$\|g\|_{1,N} := \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g\left(\frac{k}{N}\right).$$

\square

The following lemma prepares Theorem 3.3.

Lemma 3.2. For $\beta \in \mathbb{N}$, $\beta \geq 2$, let a function $g \in BV_0^{\beta-1}$ be given. Then for $N \in 2\mathbb{N}$, $N \geq 2\beta$, and $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ the following estimates hold true

$$\left| \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g\left(\frac{k}{N}\right) e^{2\pi i k x} \right| \leq \frac{(2^\beta - 1) \zeta(\beta) |g^{(\beta-1)}|_V}{(2N)^{\beta-1} |2\pi x|^\beta},$$

$$\|g\|_{1,N} \geq N \left(1 - 2\zeta(\beta) (4\pi\beta)^{-\beta} |g^{(\beta-1)}|_V\right).$$

Proof. First, we define for $x, z \in [-\frac{1}{2}, \frac{1}{2}]$ the function $h_x(z) := g(z) e^{2\pi i N x z}$. Thus, the Poisson summation formula yields

$$\frac{1}{N} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g\left(\frac{k}{N}\right) e^{2\pi i k x} = \frac{1}{N} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} h_x\left(\frac{k}{N}\right) = \sum_{r \in \mathbb{Z}} \int_{-\frac{1}{2}}^{\frac{1}{2}} h_x(z) e^{-2\pi i N r z} dz$$

and by applying integration by parts and the fact that $g^{(\gamma)}(\pm\frac{1}{2}) = 0$ for $\gamma = 0, \dots, \beta-2$ further

$$\begin{aligned} \left| \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} h_x\left(\frac{k}{N}\right) \right| &= \left| N \sum_{r \in \mathbb{Z}} (2\pi i N (r-x))^{-(\beta-1)} \int_{-\frac{1}{2}}^{\frac{1}{2}} g^{(\beta-1)}(z) e^{2\pi i N z (x-r)} dz \right| \\ &= \left| \frac{N}{(2\pi i N)^\beta} \sum_{r \in \mathbb{Z}} (r-x)^{-\beta} \int_{-\frac{1}{2}}^{\frac{1}{2}} g^{(\beta-1)}(z) \left(\frac{d}{dz} e^{2\pi i N z (x-r)}\right) dz \right| \\ &\leq \frac{1 + |x|^\beta \sum_{r \in \mathbb{Z} \setminus \{0\}} |r-x|^{-\beta}}{(2\pi)^\beta N^{\beta-1} |x|^\beta} \sup_{r_0 \in \mathbb{Z}} \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} g^{(\beta-1)}(z) \left(\frac{d}{dz} e^{2\pi i N z (x-r_0)}\right) dz \right|. \end{aligned}$$

Using $1 + |x|^\beta \sum_{r \in \mathbb{Z} \setminus \{0\}} |r-x|^{-\beta} \leq (2^\beta - 1) 2^{1-\beta} \zeta(\beta)$ for $|x| \leq \frac{1}{2}$ and

$$\left| \int_{-\frac{1}{2}}^{\frac{1}{2}} g^{(\beta-1)}(z) \left(\frac{d}{dz} e^{2\pi i N z (x-r_0)}\right) dz \right| \leq |g^{(\beta-1)}|_V$$

yields the assertion.

By the Poisson summation formula, we note furthermore that

$$\frac{1}{N} \|g\|_{1,N} \geq 1 - \left| \sum_{r \in \mathbb{Z} \setminus \{0\}} \int_{-\frac{1}{2}}^{\frac{1}{2}} g(z) e^{-2\pi i N r z} dz \right|$$

and proceed analogously in order to prove the second assertion where we use $N \geq 2\beta$ to obtain an estimate independent of N . \blacksquare

Theorem 3.3. For $\beta \in \mathbb{N}$, $\beta \geq 2$, let a function $g \in BV_0^{\beta-1}$ be given. Furthermore, let $N \in 2\mathbb{N}$, $N \geq 2\beta$, and the damping factors

$$\hat{w}_k = \frac{g\left(\frac{k}{N}\right) + g\left(\frac{k+1}{N}\right)}{2\|g\|_{1,N}}, \quad k = -\frac{N}{2}, \dots, \frac{N}{2} - 1,$$

be given. Then the kernel K_N , cf. Definition 2.1, fulfils

$$|K_N(x)| \leq \frac{(2^\beta - 1) \zeta(\beta) |g^{(\beta-1)}|_V}{2^{\beta-1} (2\pi)^\beta - \zeta(\beta) \beta^{-\beta} |g^{(\beta-1)}|_V} \frac{1}{N^\beta |x|^\beta}$$

for $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$.

Proof. Note first, that

$$K_N(x) = \frac{1 + e^{-2\pi i x}}{2\|g\|_{1,N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g\left(\frac{k}{N}\right) e^{2\pi i k x}.$$

Thus, we obtain $K_N(0) = 1$ and by applying Lemma 3.2 also the decay property. \blacksquare

We apply Theorem 3.3 in the following to the particular class of B-Spline kernels.

Definition 3.4. Let $\beta \in \mathbb{N}$ be given. The *normalised B-Spline* is defined by

$$g_\beta(z) := \beta N_\beta\left(\beta z + \frac{\beta}{2}\right),$$

where N_β denotes the cardinal B-Spline of order β . The cardinal B-Splines are given by $N_1(z) = 1$ for $z \in (0, 1)$, $N_1(z) = 0$ elsewhere, and $N_{\beta+1}(z) = \int_{z-1}^z N_\beta(\tau) d\tau$, see e.g. [4].

Furthermore, we define for $\beta \in \mathbb{N}$ and $N \in 2\mathbb{N}$ the *B-Spline kernel* by

$$B_{\beta,N}(x) := \frac{1 + e^{-2\pi i x}}{2\|g_\beta\|_{1,N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g_\beta\left(\frac{k}{N}\right) e^{2\pi i k x}.$$

\square

Corollary 3.5. Let $\beta \in \mathbb{N}$, $\beta \geq 2$, and $N \in 2\mathbb{N}$, $N \geq 2\beta$, be given. Then the B-Spline kernel $B_{\beta,N}(x)$, cf. Definition 3.4, fulfils

$$|B_{\beta,N}(x)| \leq \frac{(2^\beta - 1) \zeta(\beta) \beta^\beta}{2^{\beta-1} \pi^\beta - \zeta(\beta)} |Nx|^{-\beta}$$

for $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ and $B_{\beta,N}(0) = 1$.

Proof. Note that $g_\beta \in BV_0^{\beta-1}$. Using $N'_\beta(z) = N_{\beta-1}(z) - N_{\beta-1}(z-1)$, we conclude

$$\left|g_\beta^{(\beta-1)}\right|_V = \beta^\beta \left|N_\beta^{(\beta-1)}\right|_V = \beta^\beta \left|\sum_{\tau=0}^{\beta-1} (-1)^\tau \binom{\beta-1}{\tau} N_1(\cdot - \tau)\right|_V = (2\beta)^\beta$$

and apply Theorem 3.3. \blacksquare

Note, that in contrast to [15] the order β of the B-Spline and the degree $N \in 2\mathbb{N}$ of the kernel $B_{\beta,N}$ are independent of each other. The special case $\beta = 1$, i.e. the “top-hat function” $g_1(z) = 1$ for $|z| < \frac{1}{2}$ and $g_1(z) = 0$ elsewhere, leads to the well known *Dirichlet kernel* $B_{1,N}(x) = \frac{1}{N} \sum_{k \in I_N} e^{2\pi i k x}$. Analogously, $\beta = 2$, i.e. the “hat function” $g_2(z) = 2 - 4|z|$ for $|z| \leq \frac{1}{2}$ and $g_2(z) = 0$ elsewhere, leads to the *Fejér kernel*. The increasing localisation of the B-Spline kernels is illustrated in Figure 3.1.

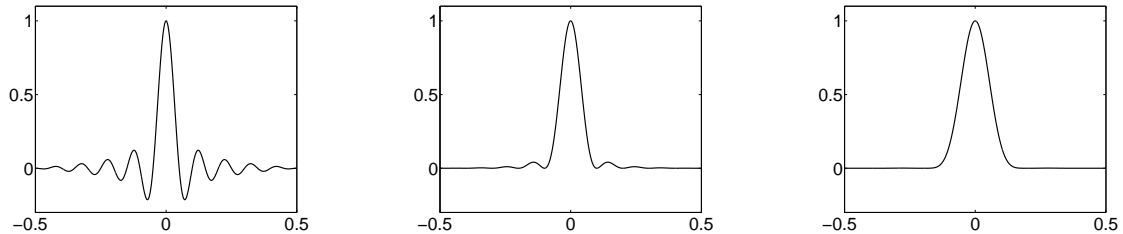


Figure 3.1: From left to right: Real part of the Dirichlet kernel $B_{1,20}$, the Fejér kernel $B_{2,20}$, and the B-Spline kernel $B_{4,20}$.

Remark 3.6. If we assume in Corollary 3.5 furthermore, that $N = \beta\sigma$, $\sigma \in \mathbb{N}$, then the stronger estimate $|B_{\beta,N}(x)| \leq 2\zeta(\beta)(1 - 2^{-\beta})(\frac{\beta}{\pi})^\beta |Nx|^{-\beta}$ holds true. This improvement is due to $\|g_\beta\|_{1,N} = N$ in Lemma 3.2 and follows from the partition of unity of the cardinal B-Spline N_β and the refinement equation $N_\beta(z) = \sum_{\tau \in \mathbb{Z}} a_\tau^{(\beta,\sigma)} N_\beta(\sigma z - \tau)$ for some finitely supported coefficients $a_\tau^{(\beta,\sigma)} > 0$, see e.g. [4, pp. 8].

In particular, the Fejér kernel fulfils $|B_{2,N}(x)| \leq |Nx|^{-2}$, $N \in 2\mathbb{N}$, which also follows from the estimate $|\sin(\pi x)| \geq 2|x|$ for $|x| \leq 1/2$ and the representation

$$B_{2,N}(x) = \frac{2(1 + e^{-2\pi i x})}{N^2} \left(\frac{\sin(\frac{N}{2}\pi x)}{\sin(\pi x)} \right)^2.$$

Along the same line follows the localisation property $B_{1,N}(x) \leq |Nx|^{-1}$ and $B_{1,N}(0) = 1$ for the Dirichlet kernel. \square

We complete this section by an extension of our result to the multivariate case $d > 1$. Indeed, tensor products of the kernels constructed in Theorem 3.3 yield also localised multivariate kernels as shown in the following corollary.

Corollary 3.7. Let the univariate kernel \tilde{K}_N , cf. Definition 2.1, fulfil for some $\beta \in \mathbb{N}$, some constant $C_\beta > 1$, and $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ the decay condition $|\tilde{K}_N(x)| \leq C_\beta |Nx|^{-\beta}$, then its tensor product kernel $K_N(\mathbf{x}) = \prod_{t=0}^{d-1} \tilde{K}_N(x_t)$ fulfils for $\mathbf{x} \in [-\frac{1}{2}, \frac{1}{2}]^d \setminus \{\mathbf{0}\}$ the estimate

$$|K_N(\mathbf{x})| \leq \frac{C_\beta}{N^\beta \|\mathbf{x}\|_\infty^\beta}.$$

Proof. The assertion follows simply from the estimate $|K_N(\mathbf{x})| \leq |\tilde{K}_N(\|\mathbf{x}\|_\infty)|$. \blacksquare

4 Stability of the interpolation problem

The nonequispaced Fourier matrix \mathbf{A} has full rank M for $d = 1$ and $N \geq M$. Unfortunately and due to the famous result of Mairhuber-Curtis, see e.g. [20, Thm. 2.3], we cannot expect full rank of the matrix \mathbf{A} for $N \geq M^{\frac{1}{d}}$ in the multivariate case $d > 1$.

However, we prove stability results for the trigonometric interpolation problem at q -separated nodes in the univariate case, cf. Theorem 4.1 and the multivariate case, cf. Theorem 4.6. Basically, a localised kernel K_N yields a diagonal dominated kernel matrix \mathbf{K}_N and thus full rank of the nonequispaced Fourier matrix \mathbf{A} for q -separated nodes. Furthermore, we prove stability results for a slightly generalised interpolation problem at equispaced nodes and subsets of equispaced nodes in Theorem 4.10. These results are applied to the B-Spline kernels from Section 3 in Corollary 4.7 and Corollary 4.11.

The univariate setting

The following Theorem 4.1 gives estimates for the extremal eigenvalues of the matrix \mathbf{K}_N under reasonable assumptions on the kernel K_N .

Theorem 4.1. *Let $N \in 2\mathbb{N}$ be given and let the kernel K_N , cf. Definition 2.1, fulfil for some $\beta > 1$ and $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ the localisation property*

$$|K_N(x)| \leq \frac{C_\beta}{N^\beta |x|^\beta}.$$

Furthermore, let a sampling set \mathcal{X} contain arbitrary nodes with separation distance $q > 0$. Then, the extremal eigenvalues of the matrix \mathbf{K}_N are bounded by

$$1 - \frac{2\zeta(\beta)C_\beta}{N^\beta q^\beta} \leq \lambda \leq 1 \leq \Lambda \leq 1 + \frac{2\zeta(\beta)C_\beta}{N^\beta q^\beta}.$$

Proof. As usual, let M denote the number of nodes in \mathcal{X} . Due to $K_N(0) = 1$, cf. Definition 2.1, we obtain $\text{trace}(\mathbf{K}_N) := \sum_{j=0}^{M-1} K_N(0) = M$. Since the trace is invariant under similarity transforms, all eigenvalues sum up to M and thus, the inequality $\lambda \leq 1 \leq \Lambda$ is fulfilled.

Now, let λ_\star be an arbitrary eigenvalue of \mathbf{K}_N , then for some index $j \in \{0, \dots, M-1\}$ Gershgorin's circle theorem yields

$$|\lambda_\star - 1| \leq \sum_{l=0; l \neq j}^{M-1} |K_N(x_j - x_l)|.$$

By using that the separation distance of the sampling set is q and by the localisation of the kernel K_N , we obtain

$$|\lambda_\star - 1| \leq \frac{C_\beta}{N^\beta} \sum_{l=0; l \neq j}^{M-1} \frac{1}{|x_j - x_l|^\beta} \leq \frac{2C_\beta}{N^\beta q^\beta} \sum_{l=1}^{\lfloor M/2 \rfloor} l^{-\beta} < \frac{2\zeta(\beta)C_\beta}{N^\beta q^\beta}. \quad \blacksquare$$

Indeed, the kernels constructed in Theorem 3.3 yield well conditioned matrices \mathbf{K}_N . We also note that the decay of the Dirichlet kernel only allows for a weaker result.

Remark 4.2. Let a sampling set $\mathcal{X} \in \mathbb{T}$ with separation distance $q > 0$ be given. Then the application of Theorem 4.1, where the last step of its proof is replaced by the estimate $\sum_{l=1}^{\lfloor M/2 \rfloor} l^{-1} \leq 1 + \ln \frac{1}{2q}$, yields: The matrix $(B_{1,N}(x_j - x_l))_{j,l=0,\dots,M-1} = \frac{1}{N} \mathbf{A} \mathbf{A}^H$ is nonsingular for $N > (1 + |\log(2q)|) q^{-1}$. The logarithmic term in this condition is clearly suboptimal. \square

As an immediate consequence of Theorem 4.1 we state a stability result for an equi-paced grid disturbed by jitter.

Corollary 4.3. Let the assumptions of Theorem 4.1 hold true. Furthermore, let the sampling nodes be of the form $x_j = -\frac{1}{2} + \frac{j - \varepsilon_j}{M}$, $j = 0, \dots, M - 1$, where $0 \leq \varepsilon_j \leq \varepsilon < 1$. Then the eigenvalues of the matrix \mathbf{K}_N are bounded by

$$1 - \frac{2\zeta(\beta) C_\beta M^\beta}{N^\beta (1 - \varepsilon)^\beta} \leq \lambda \leq 1 \leq \Lambda \leq 1 + \frac{2\zeta(\beta) C_\beta M^\beta}{N^\beta (1 - \varepsilon)^\beta}.$$

Proof. Since the separation distance is bounded by $q \geq M^{-1}(1 - \varepsilon)$ the result follows by Theorem 4.1. \blacksquare

The multivariate setting

First, we borrow a packing argument on the sphere from [15] and refine it in Lemma 4.5 for the present setting, i.e., we show how many q -separated nodes can be placed in a certain distance to a reference node, see also Figure 4.1.

Definition 4.4. For $d \in \mathbb{N}$ and a separation distance q , $0 < q \leq \frac{1}{2}$, we define the partitioning

$$R_{q,m} := \left\{ \mathbf{x} \in \mathbb{T}^d : mq \leq \text{dist}(\mathbf{x}, \mathbf{0}) < (m+1)q \right\}$$

for $m = 0, \dots, \lfloor q^{-1}/2 \rfloor - 1$ and

$$R_{q, \lfloor q^{-1}/2 \rfloor} := \left\{ \mathbf{x} \in \mathbb{T}^d : \lfloor q^{-1}/2 \rfloor q \leq \text{dist}(\mathbf{x}, \mathbf{0}) \leq 1/2 \right\}.$$

Its restriction to the sampling set \mathcal{X} will be denoted by $R_{\mathcal{X},q,m} := R_{q,m} \cap \mathcal{X}$. \square

Lemma 4.5. Let $d \in \mathbb{N}$ and an q -separated sampling set \mathcal{X} with $0 < q \leq \frac{1}{2}$ be given. Then, each of the sets $R_{\mathcal{X},q,m}$ has bounded cardinality

$$|R_{\mathcal{X},q,m}| \leq 2^d (2^d - 1) m^{d-1}, \quad m = 1, \dots, \lfloor q^{-1}/2 \rfloor.$$

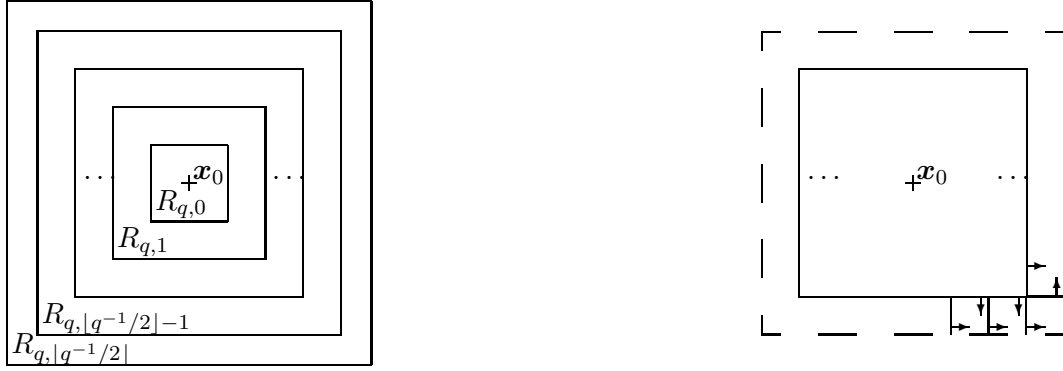


Figure 4.1: Partitioning of the torus \mathbb{T}^2 into the rings $R_{q,m}$, $m = 0, \dots, \lfloor q^{-1}/2 \rfloor$ (left). Further subdivision into shifted and rotated versions of the cube $[0, q]^d$, where arrows indicate the “ownership” of the faces to a particular cube (right).

Proof. We use a packing argument for the partition $\{R_{q,m}, m = 0, \dots, \lfloor q^{-1}/2 \rfloor\}$ of the torus \mathbb{T}^d . Each ring $R_{q,m}$ is subdivided into shifted and rotated versions of the cube $[0, q]^d$, cf. Figure 4.1 (right). This is done such that each point in $R_{q,m}$ is contained in at least one of these boxes and the boxes share no interior points with each other. Every box contains at most one node of the sampling set and hence, the estimate

$$|R_{\mathcal{X},q,m}| \leq \frac{1}{q^d} \int_{R_{q,m}} d\mathbf{x} \leq 2^d \left((m+1)^d - m^d \right) = 2^d \sum_{t=1}^d \binom{d}{t} m^{d-t} \leq 2^d m^{d-1} \sum_{t=1}^d \binom{d}{t}.$$

is valid. ■

Using localised kernels in conjunction with a separated sampling set and Lemma 4.5, we state the following theorem on the stability of the interpolation problem.

Theorem 4.6. *Let $d \in \mathbb{N}$, $N \in 2\mathbb{N}$ be given and let the kernel K_N , cf. Definition 2.1, fulfil for some $\beta > d$ and $\mathbf{x} \in [-\frac{1}{2}, \frac{1}{2}]^d \setminus \{\mathbf{0}\}$ the localisation property*

$$|K_N(\mathbf{x})| \leq \frac{C_\beta}{N^\beta \|\mathbf{x}\|_\infty^\beta}.$$

Furthermore, let a sampling set \mathcal{X} contain arbitrary nodes with separation distance $0 < q \leq \frac{1}{2}$. Then, the extremal eigenvalues of the matrix \mathbf{K}_N are bounded by

$$1 - \frac{2^d (2^d - 1) \zeta(\beta - d + 1) C_\beta}{N^\beta q^\beta} \leq \lambda \leq 1 \leq \Lambda \leq 1 + \frac{2^d (2^d - 1) \zeta(\beta - d + 1) C_\beta}{N^\beta q^\beta}.$$

Proof. Let λ_* be an arbitrary eigenvalue of \mathbf{K}_N . Without loss of generality, let the diagonal element of the matrix \mathbf{K}_N used in Gershgorin’s circle theorem correspond to

$\mathbf{x}_0 = \mathbf{0}$. Then we conclude by $K_N(\mathbf{0}) = 1$, cf. Definition 2.1, that

$$|\lambda_\star - 1| \leq \sum_{l=1}^{M-1} |K_N(\mathbf{0} - \mathbf{x}_l)|.$$

Using the partition from Definition 4.4, Lemma 4.5, and the localisation of the kernel K_N , we get

$$\begin{aligned} |\lambda_\star - 1| &\leq \sum_{m=1}^{\lfloor q^{-1/2} \rfloor} \sum_{\mathbf{x}_l \in R_{\mathcal{X}, q, m}} |K_N(-\mathbf{x}_l)| \\ &\leq \frac{2^d (2^d - 1) C_\beta}{N^\beta} \sum_{m=1}^{\lfloor q^{-1/2} \rfloor} m^{d-1} \max_{\mathbf{x} \in R_{q, m}} \|\mathbf{x}\|_\infty^{-\beta} \\ &\leq \frac{2^d (2^d - 1) \zeta(\beta - d + 1) C_\beta}{N^\beta q^\beta}. \end{aligned} \quad \blacksquare$$

Particularly, this result includes Theorem 4.1 if we set $d = 1$.

Corollary 4.7. Let the dimension $d \in \mathbb{N}$, an arbitrary sampling set $\mathcal{X} \in \mathbb{T}^d$ with separation distance $0 < q \leq \frac{1}{2}$, and a polynomial degree $N \in 2\mathbb{N}$, $N > 2dq^{-1}$, be given. Then the nonequispaced Fourier matrix \mathbf{A} has full rank. Moreover, the eigenvalues of the kernel matrix $\mathbf{K}_N = \mathbf{A} \hat{\mathbf{W}} \mathbf{A}^\mathsf{H}$ obtained from the B-Spline kernel of order $\beta = d + 1$ are bounded by

$$0 < 1 - \left(\frac{2d}{Nq}\right)^{d+1} \leq \lambda \leq 1 \leq \Lambda \leq 1 + \left(\frac{2d}{Nq}\right)^{d+1}.$$

Proof. Note first that $N \geq 2\beta$. We apply Theorem 4.6 where we use the estimates for C_β given in Corollary 3.5 and simplify the involved constant. Hence, the full rank of \mathbf{A} follows. \blacksquare

Thus, we have shown that the optimal trigonometric interpolation problem at q -separated nodes in d dimensions obeys a uniformly bounded condition number for a polynomial degree $N > 2dq^{-1}$ and appropriate damping factors. The dependence on q^{-1} is optimal as the subsequent analysis of the equispaced case shows. However, the constant $2d$ is not optimal for high spatial dimensions. As pointed out in [2] for the related approximation problem (1.2), it is an open problem to improve on this.

In summary, Lemma 2.3 and Corollary 4.7 assure in our situation a prescribed reduction of the error $\|\hat{\mathbf{e}}_l\|_{\hat{\mathbf{W}}^{-1}}$ in a constant number of iterations. Hence, if we assume an additional uniformity condition $q = cM^{-\frac{1}{d}}$ for the sampling set \mathcal{X} , the total arithmetical complexity of Algorithm 1 for solving (1.3) up to a prescribed error is bounded by $\mathcal{O}(M \log M)$.

Remark 4.8. Other frequently applied kernels also possess specific localisation properties and thus, yield stable interpolation at q -separated nodes by means of Theorem 4.6 as follows:

We define for $\beta \in 2\mathbb{N}$, $\sigma \in \mathbb{N}$, and $N = \beta(\sigma - 1) + 2$ the Jackson kernel by

$$J_{\beta,N}(x) := \frac{1 + e^{-2\pi i x}}{2\sigma^\beta} \left(\frac{\sin(\sigma\pi x)}{\sin(\pi x)} \right)^\beta.$$

Being a normalised power of the Fejér kernel $B_{2,2\sigma}$, the coefficients \hat{w}_k of the Jackson kernel can be obtained by an iterated discrete convolution of the coefficients of the Fejér kernel, see [1] for details. In contrast, the B-Spline kernel relies on a continuous convolution. The Jackson kernel is localised as $|J_{\beta,N}(x)| \leq (\frac{\beta}{2})^\beta |Nx|^{-\beta}$ for $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ and fulfils $J_{\beta,N}(0) = 1$. Hence, the tensor product Jackson kernel of appropriate order yields for $N \in 2\mathbb{N}$, $N > 2.1dq^{-1}$, the nonsingular kernel matrix

$$\left(J_{2\lceil \frac{d+1}{2} \rceil, N}(\mathbf{x}_j - \mathbf{x}_l) \right)_{j,l=0,\dots,M-1}.$$

Secondly, it is well known that the weight $1 + (2\pi k)^{2\alpha}$ is associated to the squared Sobolev norm $\|f\|_2^2 + \|f^{(\alpha)}\|_2^2$. For $\beta \in \mathbb{N}$ and $\alpha, \gamma > 0$, a regularised and slightly generalised weight is given by

$$g_{\alpha,\beta,\gamma}(z) := c_{\alpha,\beta,\gamma} \frac{(\frac{1}{4} - z^2)^\beta}{\gamma + |z|^{2\alpha}}$$

for $|z| \leq \frac{1}{2}$ and $g_{\alpha,\beta,\gamma}(z) = 0$ elsewhere, where the constant $c_{\alpha,\beta,\gamma}$ is chosen such that $\|g_{\alpha,\beta,\gamma}\|_{L_1} = 1$. Here, the denominator generalises the weight $1 + (2\pi k)^{2\alpha}$ and the nominator ensures $g_{\alpha,\beta,\gamma} \in BV_0^{\beta-1}$. We define for $N \in 2\mathbb{N}$ the Sobolev kernel by

$$S_{\alpha,\beta,\gamma,N}(x) := \frac{1 + e^{-2\pi i x}}{2\|g_{\alpha,\beta,\gamma}\|_{1,N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} g_{\alpha,\beta,\gamma}\left(\frac{k}{N}\right) e^{2\pi i k x}.$$

The kernel is localised as $|S_{\alpha,\beta,\gamma,N}(x)| \leq \tilde{c}_{\alpha,\beta,\gamma} |Nx|^{-\beta}$ for $x \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ and some constant $\tilde{c}_{\alpha,\beta,\gamma} > 0$ and fulfils $S_{\alpha,\beta,\gamma,N}(0) = 1$. \square

Results for equispaced nodes

In the case of equispaced nodes we employ the fact that the matrix \mathbf{K}_N is circulant. We present a slightly generalised result in the following Theorem 4.10.

Definition 4.9. We define for $d, n \in \mathbb{N}$ and weights $\hat{w}_{\mathbf{k}} \in \mathbb{R}$, $\mathbf{k} \in \mathbb{Z}^d$ with $\sum_{\mathbf{k} \in \mathbb{Z}^d} |\hat{w}_{\mathbf{k}}| < \infty$, the kernel

$$K(\mathbf{x}) := \sum_{\mathbf{k} \in \mathbb{Z}^d} \hat{w}_{\mathbf{k}} e^{2\pi i \mathbf{k} \mathbf{x}}$$

and by evaluating at the equispaced sampling nodes $\mathbf{j} = (j_0, \dots, j_{d-1})^\top \in I_n$, the matrix

$$\mathbf{K} := \left(K \left(\frac{\mathbf{j} - \mathbf{l}}{n} \right) \right)_{\mathbf{j}, \mathbf{l} \in I_n} \in \mathbb{C}^{n^d \times n^d}.$$

□

Theorem 4.10. *The matrix \mathbf{K} in Definition 4.9 possesses the following properties. Its eigenvalues are given by*

$$\lambda_{\mathbf{s}}(\mathbf{K}) = n^d \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{w}_{\mathbf{s} + n\mathbf{r}}$$

for $\mathbf{s} \in I_n$. For tensor product weights $\hat{w}_{\mathbf{k}} = \prod_{t=0}^{d-1} \hat{w}_{k_t}$, $\sum_{\mathbf{k} \in \mathbb{Z}^d} |\hat{w}_{\mathbf{k}}| < \infty$, this simplifies to

$$\lambda_{\mathbf{s}}(\mathbf{K}) = n^d \prod_{t=0}^{d-1} \sum_{r_t \in \mathbb{Z}} \hat{w}_{s_t + nr_t}$$

for $\mathbf{s} \in I_n$. Moreover, the extremal eigenvalues of

$$\mathbf{K}_\Gamma := \left(K \left(\frac{\mathbf{j} - \mathbf{l}}{n} \right) \right)_{\mathbf{j}, \mathbf{l} \in \Gamma}$$

are bounded by the extremal eigenvalues of \mathbf{K} for any $\Gamma \subset I_n$.

Proof. The matrix \mathbf{K} is multilevel circulant and thus diagonalised by the Fourier matrix $\mathbf{F}_n = (e^{2\pi i \mathbf{k} \mathbf{j} / n})_{\mathbf{j}, \mathbf{k} \in I_n}$. We calculate

$$\begin{aligned} \left(\mathbf{F}_n^\mathbf{H} \mathbf{K} \mathbf{F}_n \right)_{\mathbf{s}, \mathbf{t}} &= \sum_{\mathbf{j}, \mathbf{l} \in I_n} e^{-2\pi i \mathbf{s} \mathbf{j} / n} K \left(\frac{\mathbf{j} - \mathbf{l}}{n} \right) e^{2\pi i \mathbf{t} \mathbf{l} / n} \\ &= \sum_{\mathbf{k} \in \mathbb{Z}^d} \hat{w}_{\mathbf{k}} \sum_{\mathbf{j} \in I_n} e^{-2\pi i \mathbf{j} (\mathbf{s} - \mathbf{k}) / n} \sum_{\mathbf{l} \in I_n} e^{2\pi i \mathbf{l} (\mathbf{t} - \mathbf{k}) / n} \end{aligned}$$

for $\mathbf{s}, \mathbf{t} \in I_n$ and use

$$\sum_{\mathbf{j} \in I_n} e^{2\pi i \mathbf{j} (\mathbf{s} - \mathbf{k}) / n} = \begin{cases} n^d & \text{if } \frac{\mathbf{s} - \mathbf{k}}{n} \in \mathbb{Z}^d, \\ 0 & \text{otherwise.} \end{cases}$$

See also [15, Cor. 3.10, Thm. 3.11] for the univariate case.

The second assertion is due to the Kronecker product structure of the matrix \mathbf{K} in the case of tensor product kernels. The last assertion follows from the fact that removing a node is nothing else than removing its corresponding row and column in \mathbf{K} and from the interlacing property for eigenvalues, see [11, pp. 185]. ■

Now, let $d, n \in \mathbb{N}$, $N \in 2\mathbb{N}$, and the equispaced sampling set $\mathcal{X} = \frac{1}{n} I_n \subset \mathbb{T}^d$ be given. A simple consequence of Theorem 4.10 is the fact that the kernel matrix $\mathbf{K}_N = \frac{1}{N} \mathbf{A} \mathbf{A}^\mathbf{H}$ is singular whenever $N < q^{-1} = n$. Hence, the condition $N > 2dq^{-1}$ for the full rank of \mathbf{A} is optimal with respect to q . Nevertheless, we also apply Theorem 4.10 to obtain positive results in the equispaced setting for the Dirichlet and the Fejér kernel.

Corollary 4.11. Let the dimension $d \in \mathbb{N}$, a sampling set \mathcal{X} with $M \geq 2$ equispaced nodes, and a polynomial degree $N \in 2\mathbb{N}$ with $N > q^{-1} = M^{\frac{1}{d}}$ be given. Then, for the Dirichlet kernel $B_{1,N}$ and its tensor product version for $d > 1$, the extremal eigenvalues of the kernel matrix \mathbf{K}_N obey

$$\left(\frac{\lfloor Nq \rfloor}{Nq}\right)^d = \lambda \leq 1 \leq \Lambda = \left(\frac{\lceil Nq \rceil}{Nq}\right)^d.$$

Furthermore, the Fejér kernel $B_{2,N}$ yields

$$\left(1 - \frac{1}{N^2 q^2}\right)^d \leq \lambda \leq 1 \leq \Lambda \leq \left(1 + \frac{1}{N^2 q^2}\right)^d$$

where equality holds for the outmost inequalities if $N = (2\sigma + 1)q^{-1}$, $\sigma \in \mathbb{N}$. In particular, the kernel matrix \mathbf{K}_N is nonsingular.

Proof. Throughout this proof, let $n = q^{-1}$ and the damping factors be extended by $\hat{w}_{\mathbf{k}} = 0$ for $\mathbf{k} \notin I_N$. We apply for $d = 1$ the first statement of Theorem 4.10 to the weights $\hat{w}_k = \frac{1}{N}$, $k \in I_N$, of the Dirichlet kernel for $N \in 2\mathbb{N}$ and to the weights $\hat{w}_k = \frac{2}{N}(1 - \frac{|2k+1|}{N})$, $k \in I_N$, of the Fejér kernel, if $N = (2\sigma + 1)q^{-1}$, $\sigma \in \mathbb{N}$, respectively.

The assertion is little more delicate for the univariate Fejér kernel and $N \neq (2\sigma + 1)q^{-1}$. We use the representation

$$|B_{2,N}(x)| = \frac{4}{N^2} \sum_{r=0}^{\frac{N}{2}-1} \sum_{k=-r}^r e^{2\pi i k x}.$$

Now, let λ_* be an arbitrary eigenvalue of the kernel matrix \mathbf{K}_N , then Gershgorin's circle theorem yields

$$|\lambda_* - 1| \leq \sum_{l=1}^{n-1} \left| B_{2,N}\left(\frac{l}{n}\right) \right| = \frac{4n}{N^2} \sum_{r=0}^{\frac{N}{2}-1} \left(2 \left\lfloor \frac{r}{n} \right\rfloor + 1\right) - 1.$$

Since for $Q := \lfloor \frac{N-2}{2n} \rfloor$ and $R := \frac{N}{2} - 1 - nQ$ the identity

$$\sum_{r=0}^{\frac{N}{2}-1} \left\lfloor \frac{r}{n} \right\rfloor = \sum_{s=0}^{Q-1} \sum_{r=sn}^{(s+1)n-1} s + \sum_{r=nQ}^{nQ+R} Q = \frac{(N-n)^2 - (2(R+1)-n)^2}{8n}$$

holds, we proceed

$$\begin{aligned} \frac{4n}{N^2} \sum_{r=0}^{\frac{N}{2}-1} \left(2 \left\lfloor \frac{r}{n} \right\rfloor + 1\right) - 1 &= \frac{4n}{N^2} \left(2 \frac{(n-N)^2 - (2(R+1)-n)^2}{8n} + \frac{N}{2}\right) - 1 \\ &= \frac{n^2}{N^2} - \left(\frac{2(R+1)-n}{N}\right)^2 \\ &\leq \frac{n^2}{N^2}. \end{aligned}$$

The case $d > 1$ is due to the second statement in Theorem. ■

5 Numerical results

In this section, we exemplify our findings on the stability of the optimal interpolation problem (1.3) and its iterative solution by Algorithm 1.

The estimates for the condition number of the kernel matrix \mathbf{K}_N for equispaced nodes, cf. Corollary 4.11, are shown in Figure 5.1 (left). For $Nq \in \mathbb{N}$ and the Dirichlet kernel $B_{1,N}$ the matrix \mathbf{K}_N is just the identity. However, using the better localised Fejér kernel $B_{2,N}$ improves the condition number already for $N > \sqrt{3}q^{-1}$ when $Nq \notin \mathbb{N}$.

We present the effect on the stability of the interpolation problem when the equispaced nodes are perturbed by jitter error, cf. Corollary 4.3, in Figure 5.1 (right). We choose different sampling sets of size $M = 1, \dots, 100$ with equispaced nodes disturbed by 10% jitter error and evaluate the maximum condition number over 100 reruns for the Dirichlet kernel $B_{1,6M}$ and the Fejér kernel $B_{2,6M}$, respectively. The Fejér kernel produces a lower condition number which is also validated by the shown upper bound. These results confirm the theoretical results of Corollary 4.11 and Corollary 4.3.

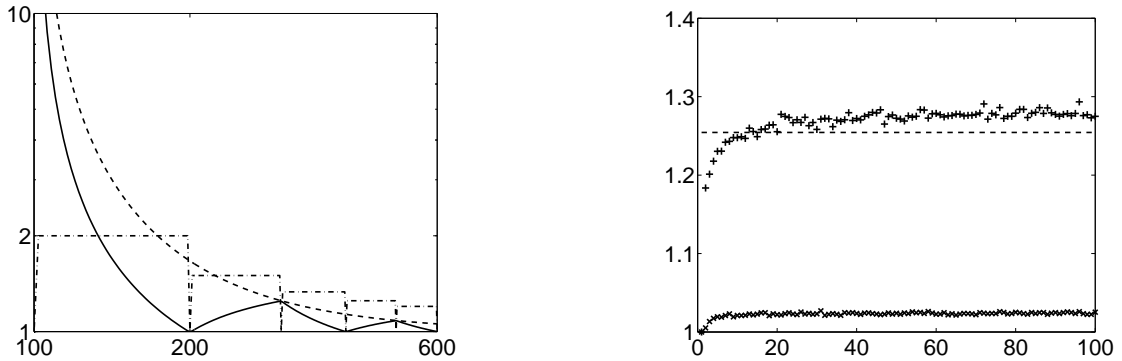


Figure 5.1: Condition number of the kernel matrix \mathbf{K}_N . Left: Condition number with respect to polynomial degree $N = 100, \dots, 600$, no weights, i.e., Dirichlet kernel (dash-dot); weight function g_2 , i.e., Fejér kernel (solid), and the estimate of Corollary 4.11 (dashed); here, the number of equispaced nodes is $M = 100$. Right: Condition number with respect to the number of nodes $M = 1, \dots, 100$, the nodes are equispaced perturbed by $\varepsilon_{\text{rel.}} = 0.1$ jitter error, the polynomial degree is $N = 6M$; no weights, i.e., Dirichlet kernel (+); weight function g_2 , i.e., Fejér kernel (\times), and its estimate by Corollary 4.3 (dashed).

Furthermore, we apply Algorithm 1 using the NFFT software package [12] to reconstruct a univariate signal from randomly scattered data in Figure 5.2 and show in Figure 5.3 the reconstruction of a bivariate signal from a glacier data set [8]. The main tool in our iterative algorithms is the NFFT, i.e., the fast matrix times vector multiplication with \mathbf{A} and \mathbf{A}^H , respectively. Details concerning NFFT algorithms can be found for example in [16] and a software package can be found in [12].

The reconstruction of the randomly sampled univariate signal shows the decay rates of our iterative scheme. The sampling set consists of $M = 100$ nodes separated by $q = 4 \times 10^{-3}$ and we reconstruct with a polynomial degree $N = 1000$ and the Dirichlet-, Fejér-, B-Spline-, and Sobolev kernel. All schemes converge within 15 iteration where this is justified only for the Fejér- and the B-Spline kernel.

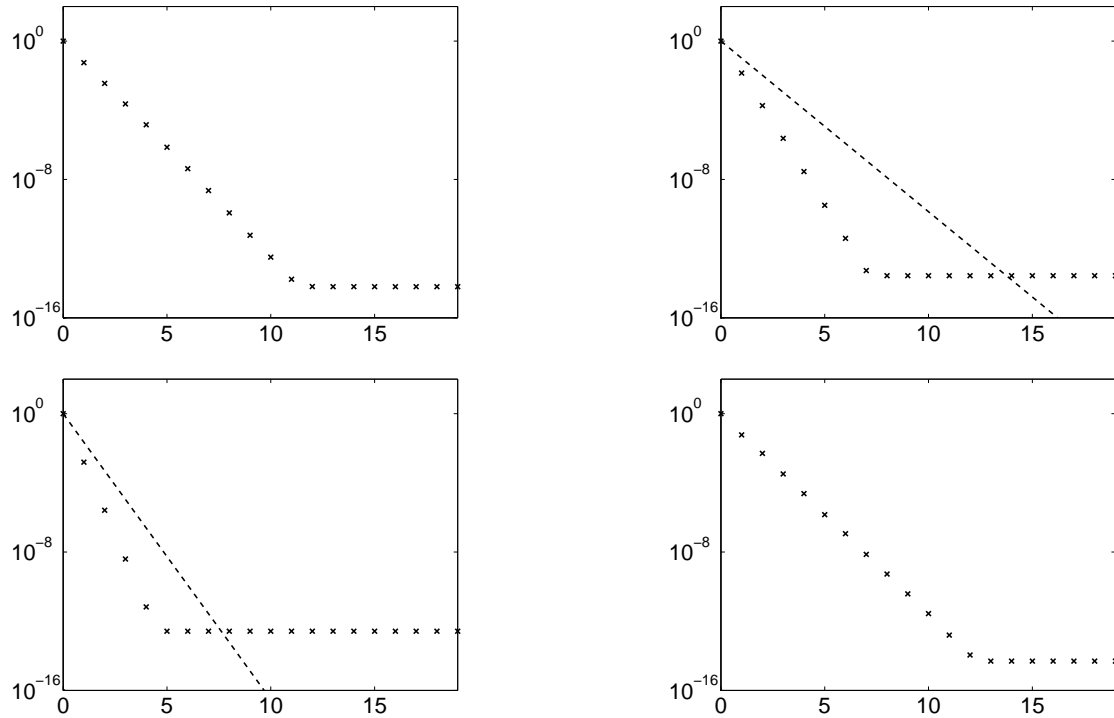


Figure 5.2: Native error $\|\hat{f}_l - \hat{W} \mathbf{A}^H \mathbf{K}_N^{-1} \mathbf{g}\|_{\hat{w}^{-1}}$ for the univariate interpolation problem with respect to the current iteration l . The number of samples is $M = 100$, the number of computed Fourier coefficients is $N = 1000$, and the separation distance of the nodes is $q = 4 \times 10^{-3}$. Top left: no weights, i.e., Dirichlet kernel; Top right: weight function g_2 , i.e., Fejér kernel, predicted decay rate (dashed); Bottom left: weight function g_4 , i.e., B-Spline kernel, predicted decay rate (dashed); Bottom right: weight function $g_{1,2,10^{-2}}$, i.e., Sobolev kernel.

The last example shows a typical test case known in radial basis function methods. We reconstruct from a data set of $M = 8345$ samples on level curves of a glacier a total number of $2^8 \times 2^8 \approx 8M$ Fourier coefficients. Note however, that the sampling set is highly nonuniform in the sense that the separation distance is very small compared to the mesh norm. The assumptions of Theorem 4.6 are not fulfilled. Nevertheless, the proposed method yields a very good approximation to the given data after 40 iterations, which is also supported by the cross validation test in Table 5.1. Here, we exclude

\tilde{M} randomly chosen samples $\tilde{\mathcal{X}} \subset \mathcal{X}$, $\mathbf{y}_{\tilde{\mathcal{X}}} \in \mathbb{C}^{\tilde{M}}$ from the reconstruction process and compare our approximations at these left out nodes. The comparison is done by means of the relative data residual and the relative validation residual after 40 iterations

$$r := \frac{\|\mathbf{y}_{\mathcal{X} \setminus \tilde{\mathcal{X}}} - \mathbf{A}_{\mathcal{X} \setminus \tilde{\mathcal{X}}} \hat{\mathbf{f}}_{40}\|_2}{\|\mathbf{y}\|_2}, \quad \tilde{r} := \frac{\|\mathbf{y}_{\tilde{\mathcal{X}}} - \mathbf{A}_{\tilde{\mathcal{X}}} \hat{\mathbf{f}}_{40}\|_2}{\|\mathbf{y}\|_2}.$$

As can be readily seen, the CGNE scheme achieves both a small data residual r and a small validation residual \tilde{r} . The proposed CGNE method combines the good data fit of the CGNR scheme ($N = 256$) with the smooth approximation of the CGNR scheme ($N = 64$).

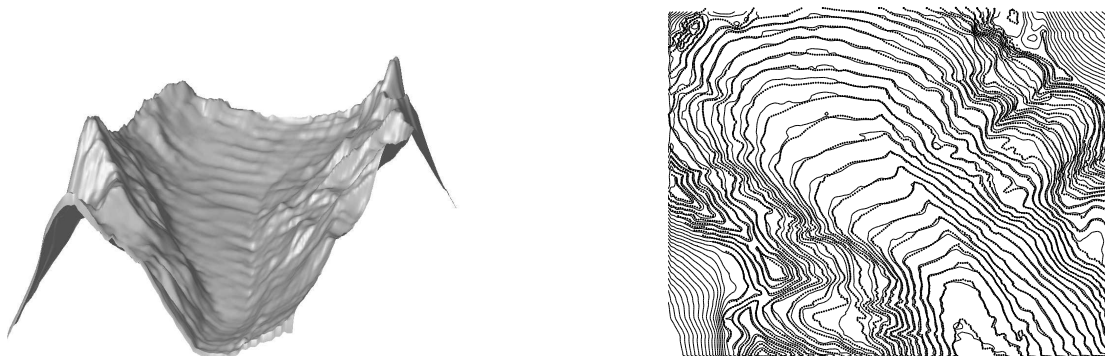


Figure 5.3: Reconstruction of the glacier data set `vol187.dat` from [8], $M = 8345$ nodes, $N = 256$, 40 iterations, tensor product damping factors $\hat{w}_{\mathbf{k}}$ to the weight function $g_{\frac{1}{2}, 3, 10^{-3}}$; see `glacier` in [12]. Left: surface plot, Right: contour plot and sampling set (\cdot).

\tilde{M}	CGNE		CGNR ($N = 256$)		CGNR ($N = 64$)	
	r	\tilde{r}	r	\tilde{r}	r	\tilde{r}
200	$6.9e - 04$	$1.7e - 02$	$5.0e - 04$	$1.4e - 01$	$8.3e - 03$	$1.7e - 02$
400	$4.7e - 04$	$2.3e - 02$	$5.0e - 04$	$2.0e - 01$	$8.3e - 03$	$2.3e - 02$
600	$5.7e - 04$	$2.9e - 02$	$5.1e - 04$	$2.5e - 01$	$8.1e - 03$	$2.9e - 02$
800	$4.7e - 04$	$3.4e - 02$	$5.0e - 04$	$2.8e - 01$	$8.0e - 03$	$3.4e - 02$
1000	$4.6e - 04$	$3.8e - 02$	$4.7e - 04$	$3.2e - 01$	$8.0e - 03$	$3.8e - 02$

Table 5.1: Cross validation of the reconstructions. The parameters of the CGNE scheme are as before. Moreover, we show the residuals of the CGNR scheme for the least squares problem (1.2) with $N = 256$ (underdetermined) and $N = 64$ (overdetermined).

6 Conclusion

We have shown that the optimal trigonometric interpolation problem at q -separated nodes in d dimensions is well conditioned for a polynomial degree $N > 2dq^{-1}$. However, in our further extensive numerical examples we observe that for $N \sim M^{1/d}$ one can expect fast convergence of Algorithm 1. If we assume furthermore a uniformity condition $q = cM^{-\frac{1}{d}}$ for the sampling set \mathcal{X} of cardinality M , then the total arithmetical complexity for solving the interpolation problem (1.3) up to a prescribed error is of order $\mathcal{O}(M \log M)$.

We remark that dependent on the application, one solves the weighted approximation problem (1.2) or the optimal interpolation problem (1.3). Under some further mild conditions on the sampling set, both problems are solved efficiently by means of the conjugate gradient method in conjunction with the nonequispaced FFT.

Acknowledgement. The first author is grateful for partial support of this work by the German Academic Exchange Service (DAAD) and the warm hospitality during his stay at the Numerical Harmonic Analysis Group, University of Vienna. We would also like to thank the referees for their valuable suggestions.

References

- [1] V. G. Alekseev. Jackson- and Jackson–Vallee Poussin-type kernels and their probability applications. *Theory Probab. Appl.*, 41:137 – 143, 1993.
- [2] R. F. Bass and K. Gröchenig. Random sampling of multivariate trigonometric polynomials. *SIAM J. Math. Anal.*, 36:773 – 795, 2004.
- [3] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.
- [4] C. K. Chui. *Multivariate Splines*. SIAM, Philadelphia, 1988.
- [5] I. Daubechies, M. Defrise, and C. D. Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57:1413 – 1457, 2004.
- [6] H. Faßbender. On numerical methods for discrete least-squares approximation by trigonometric polynomials. *Math. Comput.*, 66:719 – 741, 1997.
- [7] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numer. Math.*, 69:423 – 440, 1995.
- [8] R. Franke. <http://www.math.nps.navy.mil/~rfranke/README>.
- [9] D. Grishin and T. Strohmer. Fast multi-dimensional scattered data approximation with Neumann boundary conditions. *Lin. Alg. Appl.*, 391:99 – 123, 2004.

- [10] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comput.*, 59:181 – 194, 1992.
- [11] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
- [12] S. Kunis and D. Potts. NFFT, Softwarepackage, C subroutine library. <http://www.tu-chemnitz.de/~potts/nfft>, 2002 – 2006.
- [13] S. Kunis and H. Rauhut. Random sampling of sparse trigonometric polynomials II - orthogonal matching pursuit versus basis pursuit. Preprint 06-06, TU-Chemnitz, 2006.
- [14] H. N. Mhaskar and J. Prestin. On the detection of singularities of a periodic function. *Adv. Comput. Math.*, 12:95 – 131, 2000.
- [15] F. J. Narcowich, N. Sivakumar, and J. D. Ward. Stability results for scattered-data interpolation on euclidean spheres. *Adv. Comput. Math.*, 8:137 – 163, 1998.
- [16] D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270. Birkhäuser, Boston, 2001.
- [17] M. Rauth and T. Strohmer. Smooth approximation of potential fields from noisy scattered data. *Geophysics*, 63:85 – 94, 1998.
- [18] L. Reichel, G. S. Ammar, and W. B. Gragg. Discrete least squares approximation by trigonometric polynomials. *Math. Comput.*, 57:273 – 289, 1991.
- [19] G. Steidl. A note on fast Fourier transforms for nonequispaced grids. *Adv. Comput. Math.*, 9:337 – 353, 1998.
- [20] H. Wendland. *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2005.