
Fast Spherical/Harmonic Spline Modeling

Martin Gutting

Contents

1	Introduction and Generalized Interpolation Problems	2712
1.1	Introduction	2712
1.2	Generalized Interpolation Problems on Regular Surfaces	2713
2	Preliminaries	2715
2.1	Harmonic Splines	2715
2.2	Smoothing Splines	2720
2.3	Spherical Splines	2720
3	The Fast Multipole Method for Splines	2721
3.1	Kelvin Transform of Reproducing Kernels	2722
3.2	Adaptive Decomposition of the Domain	2723
3.3	Single Pole Expansion	2725
3.4	The Fast Multipole Algorithm	2727
3.5	Acceleration of the Translations	2728
4	Numerical Tests and Results	2733
5	Overlapping Domain Decomposition Preconditioner	2736
6	Numerical Tests and Results	2740
7	Conclusion	2742
	References	2744

Abstract

Spherical and harmonic splines are closely related approaches to solve interpolation/approximation as well as boundary value problems on the sphere and on regular (sphere-like) surfaces, respectively. In any case they lead to a system of linear equations which requires fast summation methods for the kernel sums. The fast multipole method achieves just that and is combined in this paper with a preconditioner using the same decomposition of the computational domain to

M. Gutting (✉)

Geomathematics Group, University of Siegen, Siegen, Germany

e-mail: gutting@mathematik.uni-siegen.de; martingutting@gmail.com

solve the system of linear equations resulting from spherical/harmonic splines. Due to the localizing nature of splines, regional problems can also be treated with this approach.

1 Introduction and Generalized Interpolation Problems

1.1 Introduction

Freeden (1981b) and independently Wahba (1981) developed spherical splines for interpolation and smoothing/approximation problems on the sphere, and it soon became clear that their idea can be extended to the more general case of harmonic splines introduced by Freeden (1981a). Since then, harmonic splines have been suggested for interpolation/approximation on regular surfaces as well as for the solution of boundary value problems. Convergence results have been found for both spherical splines and harmonic splines (cf. Freeden 1984a,b, 1987a,b), many different types of kernels have been introduced and even categorized (cf. Freeden and Schreiner 2014 and the references therein), and these splines became an important tool for geosciences with many applications (see, e.g., Freeden 1999; Freeden and Michel 2004; Freeden and Gerhards 2013 and the references therein).

Since splines lead to a system of linear equations which in case of harmonic splines has to be densely populated, a major obstacle has always been the solution of this system. Any iterative solver still requires fast summation methods for the splines to be efficient. On the sphere, there are several possible choices depending on the problem at hand: spherical panel clustering (cf. Freeden et al. 1998b; Freeden 1999; Fengler 2005 and the references therein), spherical FFT for gridded data points, or spherical NFFT for non-equispaced data (cf. Potts and Steidl 2003; Keiner et al. 2006).

The fast multipole method (FMM) which has been introduced first in two and then in three dimensions by Rokhlin (1985), Greengard and Rokhlin (1987, 1988), and Greengard (1988) can be combined with harmonic splines for fast summation and therefore also with spherical splines. Such a combination is used in Glockner (2002) to solve problems of satellite geodesy with harmonic splines of one specific type, namely, the singularity kernel. This kernel is also considered here, because it is the most obvious choice, but we also consider the Abel-Poisson kernel and use the accelerated version of the FMM that was first introduced in Greengard and Rokhlin (1997) and Cheng et al. (1999). This approach has also been applied to the oblique boundary value problem of potential theory in Gutting (2007, 2012).

The outline of this paper is as follows: In Sect. 1.2 the generalized interpolation problems are introduced together with the necessary basic definitions. Section 2 summarizes the theory of harmonic splines, spline smoothing as well as spherical splines. In Sect. 3 we establish the connection between harmonic splines and the sums that can be computed by the fast multipole method. We summarize the

adaptive construction of the decomposition of the computational domain and the kernel expansions that are required. We provide our version of the fast multipole algorithm for harmonic splines and briefly show the acceleration by the exponential translation introduced by Greengard and Rokhlin (1997) and Cheng et al. (1999) as well as our approach to use recombination strategies to reduce the number of translations. The parameters of the algorithm are investigated and optimized in Sect. 4. For the iterative solution of the system of linear equations, it is advisable to use some form of preconditioning. In Sect. 5 we suggest an additive Schwarz preconditioner based on an overlapping domain decomposition that is closely related to the decomposition required for the FMM. The results using the topography of the Earth as boundary surface are presented in Sect. 6, and we give a short conclusion with some outlook to the future in Sect. 7.

1.2 Generalized Interpolation Problems on Regular Surfaces

With regard to boundary value problems, we define the properties of the boundary surfaces that are considered such as the regularity of the surface.

Definition 1. A $C^{(k)}$ -regular surface $\Sigma \subset \mathbb{R}^3$ is a surface in \mathbb{R}^3 which has to fulfill the following properties:

- (i) Σ divides \mathbb{R}^3 into the interior Σ^{int} and the exterior Σ^{ext} , where Σ^{int} is a bounded region and Σ^{ext} is an unbounded region.
- (ii) The origin is contained in Σ^{int} .
- (iii) Σ is closed (and therefore compact) and free of double points.
- (iv) Σ is a $C^{(k)}$ -surface, i.e., for each $x \in \Sigma$ there exists a neighborhood $U \subset \mathbb{R}^3$ of x such that $\Sigma \cap U$ possesses a $C^{(k)}$ -parametrization.

A $C^{(k,\lambda)}$ -regular surface $\Sigma \subset \mathbb{R}^3$ with $\lambda \in (0, 1)$ is a $C^{(k)}$ -regular surface where every point $x \in \Sigma$ possesses a neighborhood U such that $\Sigma \cap U$ can locally be parameterized by a k -times λ -Hölder continuously differentiable parametrization.

The extension to $C^{(k,\lambda)}$ -regular surfaces is required for the consideration of oblique derivative boundary value problems as in Gutting (2012), we do not need it in this paper.

Now we introduce the interpolation problem on a regular surface such that interpolation on the sphere can be interpreted as a special case.

Problem 1 (Interpolation on a regular surface). Let Σ be a $C^{(0)}$ -regular surface. Let a finite set of points $\{x_1, \dots, x_N\} \subset \Sigma$ on the surface and data $F_i, i = 1, \dots, N$ corresponding to these points be given. The aim is to find a function $F \in C^{(0)}(\Sigma)$ such that $F(x_i) = F_i, i = 1, \dots, N$.

Of course, it is also possible to search for F in other function spaces if additional/other properties of the interpolating function are desired. Often the data F_i are error affected, and strict interpolation is less desirable than a good approximation that reduces the errors. Then, the interpolation conditions are reduced to $F(x_i) \approx F_i, i = 1, \dots, N$, and F has to minimize some functional that balances closeness to the data and smoothness of F , usually with one or several parameters.

The Dirichlet boundary value problem for the exterior space of a regular surface (of sufficient smoothness) in its classical form is described as follows (cf., e.g., Freeden and Gerhards 2013 and the references therein):

Problem 2. Let Σ be a $C^{(k)}$ -regular surface with $k \geq 2$. Let $F \in C^{(0)}(\Sigma)$ be a given boundary function. The task is to find a potential U with the following three properties:

- (i) $U \in C^{(0)}(\overline{\Sigma^{\text{ext}}}) \cap C^{(2)}(\Sigma^{\text{ext}})$ is harmonic in Σ^{ext} ,
- (ii) U is regular at infinity, i.e., for $|x| \rightarrow \infty$,

$$|U(x)| = \mathcal{O}(|x|^{-1}), \tag{1}$$

$$|\nabla U(x)| = \mathcal{O}(|x|^{-2}) \tag{2}$$

uniformly with respect to all directions $\frac{x}{|x|}$.

- (iii) $U^+ = F$ on Σ , i.e., for all $x \in \Sigma$,

$$U^+ = \lim_{\substack{\tau \rightarrow 0 \\ \tau > 0}} U(x + \tau \nu(x)) = F(x),$$

where ν is the unit normal vector field on Σ directed into Σ^{ext} .

It is well known that under these conditions, the Dirichlet boundary value problem possesses a unique solution (see, e.g., Freeden and Michel 2004; Freeden and Gerhards 2013). Note that the special case of $\Sigma = \Omega_r$, i.e., a sphere of radius r around the origin, is very well known, and for such results, we refer to Freeden et al. (1998a), Freeden and Michel (2004), Freeden and Gerhards (2013), and the references therein.

In this paper we focus on the discrete version of the Problem 2 which requires only the values of the boundary function in a finite set of points on the surface.

Problem 3. Let Σ be a $C^{(k)}$ -regular surface with $k \geq 2$. Let $\{x_1, \dots, x_N\} \subset \Sigma$ be a discrete set of N points on the surface. For each point x_i , let $F_i = U(x_i)$ be given, where $i = 1, \dots, N$.

The task is to determine the potential $U \in C^{(0)}(\overline{\Sigma^{\text{ext}}}) \cap C^{(2)}(\Sigma^{\text{ext}})$ which is harmonic in Σ^{ext} and regular at infinity (i.e., (1) and (2) hold) or an approximation U_N to it which fits the data, i.e., for $i = 1, \dots, N$,

$$U_N(x_i) = F_i = U(x_i).$$

This formulation reduces the Dirichlet boundary value problem to a (generalized) interpolation problem. We also want to consider regional problems on parts of the surface Σ , i.e., in Problem 3 all points $x_i \in \Gamma \subsetneq \Sigma$ where Γ denotes a region on Σ . Therefore, our approximation needs to be localizing, and we use harmonic splines as first suggested in Freedden (1981a, 1982a,b) (see also Freedden and Michel 2004; Freedden and Gerhards 2013 and the references therein for the use of harmonic splines in boundary value problems of geomathematics).

2 Preliminaries

Spherical harmonics, which we denote by $Y_{n,m}$ (with degree $n \in \mathbb{N}_0$, order $m = -n, \dots, n$), are known to form a complete orthonormal basis of the space $L^2(\Omega)$ of square-integrable functions on the unit sphere Ω (see, e.g., Edmonds 1964; Varšalovič et al. 1988; Freedden and Gutting 2013). From an approximation point of view, it is important to note that spherical harmonics $\{Y_{n,m}\}_{n \in \mathbb{N}_0, m = -n, \dots, n}$ form a closed system in $C(\Omega)$ and are closed and complete in $L^2(\Omega)$. This allows the representation of square-integrable functions on any sphere Ω_R of radius $R > 0$ by their Fourier series, where the Fourier coefficients of $F \in L^2(\Omega_R)$ are denoted by

$$F^\wedge(n, m) = \int_{\Omega_R} F(x) \frac{1}{R} Y_{n,m}\left(\frac{x}{R}\right) d\omega_R(x). \quad (3)$$

2.1 Harmonic Splines

Harmonic splines are constructed in such a way that they are subspaces of the space of harmonic functions on a sphere situated inside the Earth, the so-called Runge (or Krarup) sphere (see Moritz 2014). Due to the Runge-Walsh approximation theorem, we can use these functions which possess a larger domain of harmonicity to approximate the solution of the problem which requires harmonicity only outside the Earth's surface (see Freedden 1999; Freedden and Michel 2004 for an extensive introduction of this technique). Harmonic splines have been introduced in Freedden (1981a, 1982a,b) and Shure et al. (1982) and their convergence properties are shown in Freedden (1987a,b). They are closely related to spherical splines which we discuss in Sect. 2.3. We briefly summarize the main ideas. For more details, the reader is referred to Freedden (1999) and Freedden and Michel (2004) and the references in these two books.

At first, we define the *Runge sphere* Ω_R which is a sphere of radius R around the origin such that the exterior of the Runge sphere, i.e., Ω_R^{ext} , contains the exterior of the regular surface Σ , i.e., $\overline{\Sigma^{\text{ext}}} \subset \Omega_R^{\text{ext}}$. See also Moritz (2014) where it is called Krarup sphere.

We define Sobolev spaces of the form $\mathcal{H} = \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})$ using the Runge sphere Ω_R and a sequence $\{A_n\}_{n \in \mathbb{N}_0} \subset \mathbb{R}$ which satisfies the summability condition

$$\sum_{n=0}^{\infty} \frac{2n + 1}{4\pi A_n^2} < \infty. \tag{4}$$

Definition 2. The Sobolev space $\mathcal{H} = \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})$ is defined by

$$\mathcal{H} = \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}}) = \overline{\mathcal{E}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})}^{\|\cdot\|_{\mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})}},$$

where $\mathcal{E}(\{A_n\}; \overline{\Omega_R^{\text{ext}}}) \subset C^{(\infty)}(\overline{\Omega_R^{\text{ext}}})$ is the set of all functions that are harmonic in Ω_R^{ext} , infinitely often differentiable on the Runge sphere Ω_R and regular at infinity (i.e., (1) and (2) hold) and whose Fourier coefficients $F^\wedge(n, m)$ with respect to $L^2(\Omega_R)$ (as defined in (3)) fulfill

$$\|F\|_{\mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^2 (F^\wedge(n, m))^2 < \infty.$$

\mathcal{H} is a Hilbert space with the inner product defined by

$$\langle F, G \rangle_{\mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})} = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^2 F^\wedge(n, m) G^\wedge(n, m)$$

for $F, G \in \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})$.

It is well known (cf. Freeden 1999 and the references therein) that such a space possesses a so-called reproducing kernel (see Aronszajn 1950 for an overview on reproducing kernels in general).

Definition 3. Let U be a nonempty set and $(X, \langle \cdot, \cdot \rangle_X)$ be a separable Hilbert space of real-valued functions on U . Let $\{B_n\}_{n \in \mathbb{N}_0}$ be a complete orthonormal system in $(X, \langle \cdot, \cdot \rangle_X)$. Any function $K : U \times U \rightarrow \mathbb{R}$ of the form

$$K(x, y) = \sum_{n=0}^{\infty} K^\wedge(n) B_n(x) B_n(y) \tag{5}$$

with $x, y \in U$ and $K^\wedge(n) \in \mathbb{R}$ for $n \in \mathbb{N}_0$ is called an X -product kernel (briefly an X -kernel).

An X -kernel $K(\cdot, \cdot) : U \times U \rightarrow \mathbb{R}$ is called a *reproducing kernel* (or shortly *repro-kernel*) for $(X, \langle \cdot, \cdot \rangle_X)$ if:

- (i) $K(x, \cdot) \in X$ for all $x \in U$.
- (ii) $\langle K(x, \cdot), F \rangle_X = F(x)$ for all $x \in U$ and all $F \in X$.

If there exists such a repro-kernel in X , then X is called a reproducing kernel Hilbert space.

It is well known that a reproducing kernel is always unique, and its existence is equivalent to the boundedness of all evaluation functionals (cf. Aronszajn 1950). In the space $\mathcal{H} = \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})$ with a summable sequence $\{A_n\}$, the repro-kernel (5) can be represented by its expansion in Legendre polynomials due to the addition theorem for spherical harmonics:

$$K_{\mathcal{H}}(x, y) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi A_n^2} \frac{1}{|x||y|} \left(\frac{R^2}{|x||y|} \right)^n P_n \left(\frac{x}{|x|} \cdot \frac{y}{|y|} \right). \tag{6}$$

We use these reproducing kernels to define harmonic splines.

Definition 4. Let $\{\mathcal{L}_1, \dots, \mathcal{L}_N\} \subset \mathcal{H}^*$ be a set of N linearly independent bounded linear functionals on the reproducing kernel Hilbert space \mathcal{H} .

Then any function S of the form

$$S = \sum_{i=1}^N a_i \mathcal{L}_i K_{\mathcal{H}}(\cdot, \cdot)$$

with a set of so-called *spline coefficients* $\{a_1, \dots, a_N\} \subset \mathbb{R}$ is called an \mathcal{H} -*spline* relative to $\{\mathcal{L}_1, \dots, \mathcal{L}_N\}$. The function space of all \mathcal{H} -splines relative to $\{\mathcal{L}_1, \dots, \mathcal{L}_N\}$ is denoted by $\mathcal{S}_{\mathcal{H}}(\mathcal{L}_1, \dots, \mathcal{L}_N)$.

It should be noted that by construction, any \mathcal{H} -spline is harmonic. The interpolating spline S^F for the function $F \in \mathcal{H}$ has to fulfill the interpolation conditions

$$\mathcal{L}_i S^F = \mathcal{L}_i F \quad \text{for } i = 1, \dots, N. \tag{7}$$

The interpolation conditions (7) can be rewritten as a system of linear equations for the spline coefficients a_j :

$$\sum_{i=1}^N a_i \mathcal{L}_i \mathcal{L}_j K_{\mathcal{H}}(\cdot, \cdot) = \mathcal{L}_j F, \quad j = 1, \dots, N, \tag{8}$$

whose corresponding matrix possesses the entries $\mathcal{L}_i \mathcal{L}_j K_{\mathcal{H}}(\cdot, \cdot)$ and is symmetric and positive definite (for linear functionals $\mathcal{L}_1, \dots, \mathcal{L}_N \in \mathcal{H}^*$ which are linearly independent).

Note that in this paper, we consider only evaluation functionals \mathcal{L}_x , i.e., $\mathcal{L}_x F = F(x)$ where $x \in \overline{\Sigma}^{\text{ext}}$. Furthermore, $\mathcal{L}_1, \dots, \mathcal{L}_N$ are given by $\mathcal{L}_i F = F(x_i)$ where $x_i \in \Sigma$. Obviously, this suffices to treat Problem 3 as well as any interpolation problems. The concept of splines can be used for a more general class of problems. In the following theorem, we summarize the properties of \mathcal{H} -splines.

Theorem 1. *Let $F \in \mathcal{H}$ and let $\{\mathcal{L}_1, \dots, \mathcal{L}_N\} \subset \mathcal{H}^*$. Then the \mathcal{H} -spline interpolation problem with the interpolation conditions (7) is uniquely solvable, and its solution $S^F \in \mathcal{S}_{\mathcal{H}}(\mathcal{L}_1, \dots, \mathcal{L}_N)$ possesses the following properties:*

- (i) S^F is the \mathcal{H} -orthogonal projection of F onto $\mathcal{S}_{\mathcal{H}}(\mathcal{L}_1, \dots, \mathcal{L}_N)$.
- (ii) $\|S^F\|_{\mathcal{H}} \leq \|F\|_{\mathcal{H}}$.
- (iii) If $G \in \mathcal{H}$ also satisfies the interpolation conditions (7), then the first minimum property holds:

$$\|G\|_{\mathcal{H}}^2 = \|S^F\|_{\mathcal{H}}^2 + \|G - S^F\|_{\mathcal{H}}^2,$$

i.e., S^F is the interpolating function of F in \mathcal{H} with minimal norm.

- (iv) If $G \in \mathcal{H}$ also satisfies (7) and $S \in \mathcal{S}_{\mathcal{H}}(\mathcal{L}_1, \dots, \mathcal{L}_N)$, then the second minimum property holds:

$$\|S - G\|_{\mathcal{H}}^2 = \|S^F - G\|_{\mathcal{H}}^2 + \|S - S^F\|_{\mathcal{H}}^2.$$

For the proof and for further details on splines, the reader is referred to Freeden (1981a, 1987a, 1999) and Freeden et al. (1998a) and the references therein.

The choice $A_n = h^{-\frac{n}{2}}$, $h \in (0, 1)$, fulfills (4) and provides us with the reproducing kernel called *Abel-Poisson kernel* (see Fig. 1) which is given by

$$K_{\mathcal{H}}(x, y) = \frac{1}{4\pi} \frac{|x|^2|y|^2 - h^2 R^4}{(|x|^2|y|^2 + h^2 R^4 - 2hR^2 x \cdot y)^{\frac{3}{2}}}. \tag{9}$$

The sequence $A_n = (n + \frac{1}{2})^{\frac{1}{2}} h^{-\frac{n}{2}}$, $h \in (0, 1)$, also satisfies (4) and leads to the *singularity kernel* (see Fig. 2) given by

$$K_{\mathcal{H}}(x, y) = \frac{1}{2\pi} \frac{1}{(|x|^2|y|^2 + h^2 R^4 - 2hR^2 x \cdot y)^{\frac{1}{2}}}. \tag{10}$$

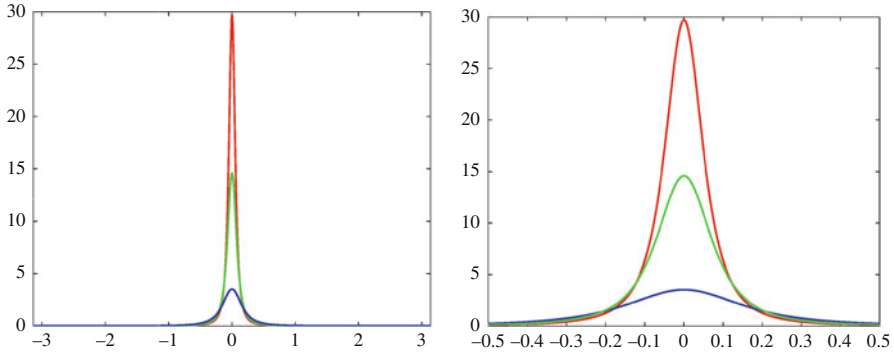


Fig. 1 The profiles of the Abel-Poisson kernel with $h = 0.93$ (red), $h = 0.9$ (green), $h = 0.8$ (blue), $|x| = |y| = 1$, $R = 0.999$, the angle between $\frac{x}{|x|}$ and $\frac{y}{|y|}$ running from $-\pi$ to π (right: a more detailed view)

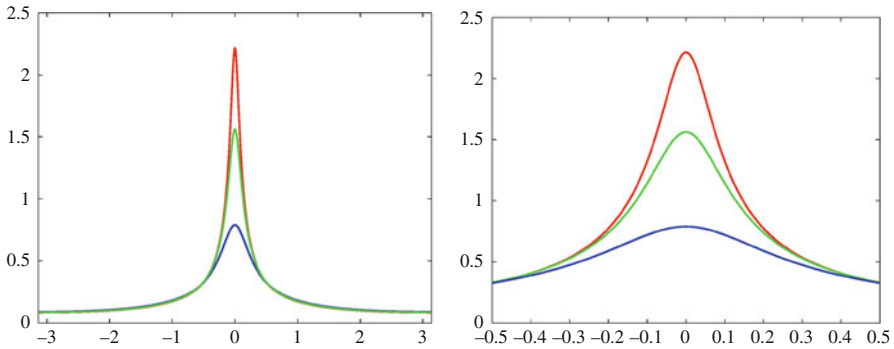


Fig. 2 The profile of the singularity kernel with $h = 0.93$ (red), $h = 0.9$ (green), $h = 0.8$ (blue), $|x| = |y| = 1$, $R = 0.999$, the angle between $\frac{x}{|x|}$ and $\frac{y}{|y|}$ running from $-\pi$ to π (right: a more detailed view)

By the Runge-Walsh approximation theorem and an extension of Helly’s theorem (cf. Yamabe 1950), it is possible to prove the existence of approximations fulfilling an interpolation condition. Moreover, convergence results for harmonic splines can be derived that show the convergence to the solution of the continuous boundary value problem (Problem 2) for an increasing density of data points, i.e., if the largest data gap goes to zero (cf. Freeden 1987a). For a detailed analysis, the reader is referred to Freeden and Michel (2004) and the literature therein.

In this paper we focus on two specific types of splines (using Abel-Poisson and singularity kernels) and propose methods to quickly compute the sums $\sum a_i K_{\mathcal{H}}(x_i, y_j)$ for many points. This can be used to solve the systems of linear equations (8) that occur in the solution of the interpolation problems using harmonic splines.

2.2 Smoothing Splines

For noisy data, i.e., $F_i = U(x_i) + \varepsilon_i \approx U(x_i)$, $i = 1, \dots, N$, in Problem 3, it makes no sense to compute an interpolation problem, but an approximation to U which smoothes the data is required (see Freeden 1981b; Wahba 1990; Freeden et al. 1997, 1998a for the case of spherical splines, Freeden (1981a, 1999) for the case of harmonic splines). This smoothing process is achieved by minimizing the following functional

$$\mu(S) = \sum_{i=1}^N \sum_{j=1}^N (\mathcal{L}_i S - F_i) \mathbf{C}_{ij} (\mathcal{L}_j S - F_j) + \beta \|S\|_{\mathcal{H}}$$

in the reproducing kernel Hilbert space $\mathcal{H} = \mathcal{H}(\{A_n\}; \overline{\Omega_R^{\text{ext}}})$. $\mathbf{C} = (\mathbf{C}_{ik}) \in \mathbb{R}^{N \times N}$ denotes a positive definite matrix, and $\beta > 0$ is a constant. The following theorem of Freeden (1999) summarizes the existence and uniqueness of smoothing splines.

Theorem 2. *Let F_i , $i = 1, \dots, N$, correspond to a set of linearly independent bounded linear functionals $\mathcal{L}_1, \dots, \mathcal{L}_N \in \mathcal{H}^*$.*

Then there exists a unique element $S \in \mathcal{S}_{\mathcal{H}}(\mathcal{L}_1, \dots, \mathcal{L}_N)$ such that $\mu(S) \leq \mu(F)$ for all $F \in \mathcal{H}$ and $\mu(S) = \mu(F)$ if and only if $S = F$. This element is called the smoothing spline. Its spline coefficients a_i , $i = 1, \dots, N$ are uniquely determined by the system of linear equations

$$\sum_{i=1}^N a_i (\mathcal{L}_i \mathcal{L}_j K_{\mathcal{H}}(\cdot, \cdot) + \beta (\mathbf{C}^{-1})_{ij}) = \mathcal{L}_j F, \quad j = 1, \dots, N. \quad (11)$$

The matrix in (11) corresponds to the one in (8) plus $\beta \mathbf{C}^{-1}$ and is still positive definite. If \mathbf{C} is the unit matrix, there is only the one smoothing parameter β , using a diagonal matrix as \mathbf{C} it is possible to introduce weights for the data F_i . The general case allows the use of \mathbf{C} to include covariance information on the data.

The choice of the smoothing parameter(s) can be interpreted as the application of a parameter choice method in the regularization theory of ill-posed problems. We refer to Bauer and Lukas (2011) and Bauer et al. (2014) for an up-to-date overview of methods and literature in the field.

2.3 Spherical Splines

Spherical splines have been introduced by Freeden (1981b) and independently by Wahba (1981) and can be embedded naturally in the harmonic setting of Sect. 2.1. On the sphere, i.e., $\Sigma = \Omega_R$, no additional Runge sphere is required and the surface itself can be used to consider the corresponding Sobolev spaces $\mathcal{H} = \mathcal{H}(\{A_n\}; \Omega_R)$

with respect to a sequence $\{A_n\}$. The summability condition (4) remains the same, and if it is satisfied by $\{A_n\}$, we again obtain a reproducing kernel. It should be remarked that on the sphere, this reproducing kernel is a radial basis function depending only on the distance $|x - y|$ of its two arguments $x, y \in \Omega_R$. Moreover, it is possible to consider Sobolev spaces that possess a locally supported reproducing kernel, i.e., $K_{\mathcal{H}}(x, y) = 0$ if $|x - y|$ is larger than a threshold value (see Freeden et al. 1998a and also the categorization in Freeden and Schreiner 2014).

For our purposes, the problem of needing a fast method to evaluate sums $\sum a_i K_{\mathcal{H}}(x_i, y_j)$ and to solve the system (8) for the spline coefficients remains the same as before. However, there exists a wider spectrum of fast summation methods on the sphere such as panel clustering (see Freeden et al. 1998b; Freeden 1999; Fengler 2005 and the references therein), spherical FFT for gridded data points or spherical NFFT for non-equispaced data (cf., e.g., Potts and Steidl 2003; Keiner et al. 2006).

In Freeden et al. (1998a), a more general construction which is orthogonal to the spherical harmonics of degrees 0 to m is introduced which can be used for approximation by combining it with the corresponding spherical harmonics. A combined interpolation and smoothing approach is included as well as spline exact numerical integration formulas on the sphere.

Convergence results for spherical splines are shown in Freeden (1984a,b) if the size of the largest data gap goes to zero. We refer the reader to Freeden et al. (1998a) and Michel (2013) and the references therein for a detailed introduction. It should be noted that the construction has also been carried over to the three-dimensional ball (see Michel 2013, 2014b and the references therein).

3 The Fast Multipole Method for Splines

In case of harmonic splines, the interpolation conditions lead to a system of linear equations with a dense matrix whose size is the number of data points. Thus, the matrix can be large, and the solution of the corresponding system of linear equations becomes difficult. For spherical splines, we are confronted with the same situation if we restrict ourselves to the Abel-Poisson kernel and the singularity kernel. Other (locally supported) spherical kernels can reduce this problem to some degree by leading to band matrices.

Each reproducing kernel corresponding to a space \mathcal{H} defined by the summable sequence $\{A_n\}$ can be written as an infinite Legendre expansion as in (6). For certain special cases such as the singularity kernel (10) and the Abel-Poisson kernel (9), the expansion can be reduced to an elementary function. These two reproducing kernels are closely related to the single pole $\frac{1}{|x-y|}$. Because of this connection, the fast multipole method (FMM), which has been introduced by Rokhlin (1985), Greengard and Rokhlin (1987, 1988), and Greengard (1988), allows the fast summation of harmonic splines, i.e., of the sum $\sum a_i K_{\mathcal{H}}(x_i, \cdot)$. Thus, the FMM also accelerates the fast computation of the matrix-vector products occurring in an iterative solver for (8).

The two main ideas of the FMM can be described as:

- Subdivision of the computational domain into hierarchical sets of nested cubes that are organized in an octree data structure,
- Interaction of cubes instead of single points by summarizing all information of a cube in coefficients of inner/outer harmonics expansions.

This means that cubes interact in two ways: directly, i.e., the kernel is evaluated for all points in one cube combined with all points in another cube, and the approximate way, i.e., by translation of the expansion coefficients. The key is to apply the approximation as often as possible and on the coarsest possible level of the tree data structure of cubes. Direct evaluation is used only for the closest cubes where the approximation is not applicable.

The FMM has seen various updates that increase its efficiency (cf., e.g., White and Head-Gordon 1996; Greengard and Rokhlin 1997; Cheng et al. 1999). We summarize our implementation and show the application of the FMM to harmonic and spherical splines. For a more detailed analysis, the reader is referred to Gutting (2007).

3.1 Kelvin Transform of Reproducing Kernels

The combination between our reproducing kernels and the fundamental solution of the Laplace equation is the Kelvin transform which can be interpreted as a reflection on a sphere of radius R around the origin. It is well known from textbooks on potential theory (cf., e.g., Kellogg 1967; Freeden and Gerhards 2013).

Definition 5. Let $\Gamma \subseteq \mathbb{R}^3$ be a domain, $W : \Gamma \rightarrow \mathbb{R}$ a function. Let the reflection of Γ on the sphere Ω_R be given by

$$\Gamma^{\text{KT}} = \left\{ x^{\text{KT}} \in \mathbb{R}^3 : \frac{R^2}{|x^{\text{KT}}|^2} x^{\text{KT}} = x \in \Gamma \right\}.$$

The function

$$W^{\text{KT}} : \Gamma^{\text{KT}} \rightarrow \mathbb{R},$$

$$x^{\text{KT}} \mapsto W^{\text{KT}}(x^{\text{KT}}) = \frac{R}{|x^{\text{KT}}|} W\left(\frac{R^2}{|x^{\text{KT}}|^2} x^{\text{KT}}\right) = \frac{R}{|x^{\text{KT}}|} W(x),$$

is called the *Kelvin transform* of W with respect to the sphere of radius R .

The Kelvin transform is applied to the reproducing kernels with respect to one argument (the other is kept fixed). The Kelvin transform $K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}})$ of the singularity kernel (10) can be computed by its expansion

$$\begin{aligned}
K_{\mathcal{H}}(x, y) &= \sum_{n=0}^{\infty} \frac{h^n}{n + \frac{1}{2}} \frac{2n + 1}{4\pi |x||y|} \left(\frac{R^2}{|x||y|} \right)^n P_n \left(\frac{x}{|x|} \cdot \frac{y}{|y|} \right) \\
&= \frac{1}{2\pi |y|} \sum_{n=0}^{\infty} \frac{(h|y^{\text{KT}}|)^n}{|x|^{n+1}} P_n \left(\frac{x}{|x|} \cdot \frac{y^{\text{KT}}}{|y^{\text{KT}}|} \right) \\
&= \frac{1}{2\pi |y|} \frac{1}{|x - hy^{\text{KT}}|} = \frac{|y^{\text{KT}}|}{R} K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}}),
\end{aligned} \tag{12}$$

where $y^{\text{KT}} = \frac{R^2}{|y|^2} y$ and

$$K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}}) = \frac{1}{2\pi R} \frac{1}{|x - hy^{\text{KT}}|}. \tag{13}$$

The Kelvin transform $K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}})$ of the Abel-Poisson kernel (9) is given by

$$\begin{aligned}
K_{\mathcal{H}}(x, y) &= \frac{1}{4\pi} \frac{|x|^2|y|^2 - h^2R^4}{(|x|^2|y|^2 + h^2R^4 - 2hR^2x \cdot y)^{\frac{3}{2}}} \\
&= \frac{|y^{\text{KT}}|}{R} \frac{1}{4\pi R} \frac{|x|^2 - h^2|y^{\text{KT}}|^2}{|x - hy^{\text{KT}}|^3} = \frac{|y^{\text{KT}}|}{R} K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}}),
\end{aligned}$$

which is related to (13) by

$$K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}}) = \frac{1}{2\pi R} (-x \cdot \nabla_x - \frac{1}{2}\text{Id}) \frac{1}{|x - hy^{\text{KT}}|}. \tag{14}$$

We can summarize both (13) and (14) by the use of the operator D_x such that

$$K_{\mathcal{H}}^{\text{KT}}(x, y^{\text{KT}}) = \frac{1}{2\pi R} D_x \frac{1}{|x - hy^{\text{KT}}|}, \tag{15}$$

where $D_x = \text{Id}$ (singularity kernel) or $D_x = -x \cdot \nabla_x - \frac{1}{2}\text{Id}$ (Abel-Poisson kernel).

3.2 Adaptive Decomposition of the Domain

At first, a bounding cube is determined which is large enough such that it contains all points x_i and targets $hy_j^{\text{KT}} = h \frac{R^2}{|y_j|^2} y_j$. This single bounding cube forms level 0 of the octree structure and is subdivided into eight equally sized cubes of half its edge length which then form level 1. Each cube is adaptively divided into nested cubes where a cube of level l has half the edge length of a cube of level $l - 1$ as proposed, e.g., by Cheng et al. (1999). The adaptive splitting is achieved by sorting the points and targets into the currently available cubes. If a

cube contains more than the prescribed maximal number of points or targets m , it is split into eight cubes of the next level, and all its points/targets are redistributed into these eight cubes. All the cubes carry information about other surrounding cubes (their so-called neighbors) which is required for the application of the kernel expansion.

The necessary vocabulary (see also Greengard and Rokhlin 1997; Cheng et al. 1999) is summarized in the following definitions.

Definition 6.

- A cube C is called *child* of the cube B if C results from a single subdivision of B which in return is named the *parent* of C .
- A cube that is not further subdivided is called *childless* or a *leaf*.
- Cubes are said to be *neighbors* if they are of the same size (same level) and share at least one boundary point. Each cube is a neighbor of itself.
- If two cubes are at the same level, but are no neighbors, they are called *well separated*, i.e., between these cubes exists at least one cube of their size.

Each cube carries the relevant information about other cubes in four lists as suggested by Cheng et al. (1999).

Definition 7.

- In *list 1* of the childless cube X are all childless cubes directly adjacent to X . List 1 only contains any cubes if X is a leaf, then it always contains at least X itself.
- *List 2* of a cube X consists of all children of neighbors of the parent cube of X which are well separated from X . The cube X does not need to be childless.
- Children of neighbors of the leaf X (or smaller cubes descending from neighbors of X) which do not have any point in common with X form *list 3*. Their parents have to be adjacent to X . If X is not childless, then list 3 is empty.
- *List 4* consists of childless cubes which are neighbors of the parent cube of X , but these childless cubes are not adjacent to X .

Notice the following observations:

- (i) List 1 is the list of all neighbors.
- (ii) All cubes in list 2 of a cube X are of the same size as X .
- (iii) The elements of list 3 are all smaller than X , and the distance between them and X is at least their side length and at most the side length of X .
- (iv) List 4 of a cube X only contains cubes that are larger than X . They are separated from X by a distance that is at least the length of X and at most their own edge length.
- (v) All members of list 1 and list 4 are leaves, and list 1 and list 3 of a cube X remain empty if X is not childless.

After finishing the construction of the octree and sorting all points and targets into cubes, the algorithm removes childless cubes that contain neither points nor targets.

3.3 Single Pole Expansion

In addition to the decomposition of the domain, the other part of the FMM is the expansion of the single pole. This is achieved similar to (12) by

$$\begin{aligned} \frac{1}{|x-y|} &= \frac{1}{|y-x_0-(x-x_0)|} \\ &= \sum_{n=0}^{\infty} \frac{|x-x_0|^n}{|y-x_0|^{n+1}} P_n \left(\frac{y-x_0}{|y-x_0|} \cdot \frac{x-x_0}{|x-x_0|} \right) \\ &= \sum_{n=0}^{\infty} \sum_{m=-n}^n I_{n,m}^*(x-x_0) O_{n,m}(y-x_0), \end{aligned} \quad (16)$$

where $|y-x_0| > |x-x_0|$ for the expansion center $x_0 \in \mathbb{R}^3$. The upper star * in (16) denotes the complex conjugate. Thereby, we use the (complex-valued) outer and inner harmonics for $n \in \mathbb{N}_0, m = -n, \dots, n$:

$$\begin{aligned} O_{n,m}(x) &= \sqrt{\frac{4\pi}{2n+1}} \frac{\sqrt{(n+m)!(n-m)!}}{|x|^{n+1}} Y_{n,m} \left(\frac{x}{|x|} \right) \\ &= (-1)^m \frac{(n-m)!}{|x|^{n+1}} P_{n,m}(\cos \vartheta) e^{im\varphi}, \\ I_{n,m}(x) &= \sqrt{\frac{4\pi}{2n+1}} \frac{|x|^n}{\sqrt{(n+m)!(n-m)!}} Y_{n,m} \left(\frac{x}{|x|} \right) \\ &= (-1)^m \frac{|x|^n}{(n+m)!} P_{n,m}(\cos \vartheta) e^{im\varphi}, \end{aligned}$$

where ϑ, φ are the spherical coordinates of $\frac{x}{|x|}$. Note that $P_{n,m}$ are the associated Legendre functions

$$P_{n,m}(t) = (1-t^2)^{\frac{m}{2}} \frac{d^m}{dt^m} P_n(t), \quad t \in [-1, 1], \quad n \geq m \geq 0.$$

The symmetry relation $P_{n,-m}(t) = (-1)^m \frac{(n-m)!}{(n+m)!} P_{n,m}(t)$ extends them for negative orders (cf., e.g., Edmonds 1964; Biedenharn and Louck 1981).

It is well known that translation theorems for these outer and inner harmonics allow to shift the expansion center (see, e.g., Epton and Dembart 1995 for a detailed derivation).

Theorem 3 (Translation Theorem for Outer Harmonics). *Let $x, y \in \mathbb{R}^3$ such that $|x| > |y|$. Then the outer harmonic of degree $n \in \mathbb{N}_0$ and order $m \in \mathbb{Z}$, $-n \leq m \leq n$, at $x - y$ can be expanded in terms of inner and outer harmonics as follows:*

$$\begin{aligned} O_{n,m}(x - y) &= \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} I_{n',m'}^*(y) O_{n+n',m+m'}(x) \\ &= \sum_{n'=n}^{\infty} \sum_{m'=-n'}^{n'} I_{n'-n,m'-m}^*(y) O_{n',m'}(x). \end{aligned} \tag{17}$$

Note that in (17), we make use of the convention that $I_{n,m} = 0$ if $|m| > n$.

Theorem 4 (Translation Theorem for Inner Harmonics). *Let $x, y \in \mathbb{R}^3$. Then the inner harmonic of degree $n \in \mathbb{N}_0$ and order $m \in \mathbb{Z}$, $-n \leq m \leq n$, at $x - y$ can be expanded in a finite sum of inner harmonics*

$$I_{n,m}(x - y) = \sum_{n'=0}^n \sum_{m'=-n'}^{n'} (-1)^{n'} I_{n',m'}(y) I_{n-n',m-m'}(x).$$

For orders with $|m| > n$, we have again by convention $I_{n,m} = 0$.

By applying (17) Theorem 3 allows the translation of an outer harmonics expansion with expansion center x_0 such as

$$F(x) = \sum_{n=0}^{\infty} \sum_{m=-n}^n F_{x_0}^{\wedge, O}(n, m) O_{n,m}(x - x_0) \tag{18}$$

which converges uniformly for $x \in \Omega_{r_0}^{\text{ext}}(x_0)$ with some $r_0 > 0$. $\Omega_{r_0}^{\text{ext}}(x_0)$ denotes the exterior of the sphere of radius r_0 around x_0 . The outer harmonics series resulting from the translation possesses the expansion center x_1 and the coefficients

$$F_{x_1}^{\wedge, O}(n', m') = \sum_{n=0}^{n'} \sum_{m=-n}^n F_{x_0}^{\wedge, O}(n, m) I_{n'-n, m'-m}^*(x_0 - x_1). \tag{19}$$

The expansion converges uniformly for $x \in \Omega_{r_1}^{\text{ext}}(x_1)$ where $\Omega_{r_1}^{\text{ext}}(x_1) \subset \Omega_{r_0}^{\text{ext}}(x_0)$. This translation is called *multipole-to-multipole translation (M2M)*.

By Theorem 3, we also find that the outer harmonics expansion can be translated into an inner harmonics series centered around x_2 which converges uniformly for $x \in \Omega_{r_2}^{\text{int}}(x_2)$ if the new ball of convergence is situated completely in $\Omega_{r_1}^{\text{ext}}(x_1)$, i.e., $\overline{\Omega_{r_1}^{\text{int}}(x_1)} \cap \overline{\Omega_{r_2}^{\text{int}}(x_2)} = \emptyset$. The resulting coefficients of the inner harmonics expansion are

$$F_{x_2}^{\wedge, I}(n', m') = \sum_{n=0}^{\infty} \sum_{m=-n}^n F_{x_1}^{\wedge, O}(n, m) (-1)^{n'+m} O_{n+n', m'-m}^*(x_2 - x_1) \quad (20)$$

and this translation is named *multipole-to-local translation (M2L)*.

Furthermore, Theorem 4 lets us shift the expansion center of such inner harmonics expansions to the new center x_3 which possesses the coefficients

$$F_{x_3}^{\wedge, I}(n', m') = \sum_{n=n'}^{\infty} \sum_{m=-n}^n F_{x_2}^{\wedge, I}(n, m) I_{n-n', m-m'}(x_3 - x_2). \quad (21)$$

and converges uniformly for $x \in \Omega_{r_3}^{\text{int}}(x_3) \subset \Omega_{r_2}^{\text{int}}(x_2)$. This translation step is called *local-to-local translation (L2L)*. For further details, we refer to Gutting (2007) and the references therein, in particular Epton and Dembart (1995).

3.4 The Fast Multipole Algorithm

Before kernel expansions can be translated, we have to compute multipole expansions, more precisely a first set of expansion coefficients for each cube containing any points. Thus, only the part of the spline related to a single cube X is considered, i.e., the kernel functions $K_{\mathcal{H}}(x_i, \cdot)$, where $x_i \in X$:

$$F = \sum_{\substack{i=1 \\ x_i \in X}}^N a_i K_{\mathcal{H}}(x_i, \cdot) = \sum_{\substack{i=1 \\ x_i \in X}}^N a_i \left(\frac{|y^{\text{KT}}|}{R} \frac{1}{2\pi R} D_x \frac{1}{|x - hy^{\text{KT}}|} \right) \Bigg|_{x=x_i}.$$

We find the following expansion for $|hy^{\text{KT}} - x_0| > |x_i - x_0|$, $x_i \in X$, i.e., if x_0 is the center of the cube X , the targets hy^{KT} and the cube X need to fulfill a distance requirement.

$$\begin{aligned} F &= \frac{|y^{\text{KT}}|}{R} \sum_{\substack{i=1 \\ x_i \in X}}^N a_i \left(\frac{1}{2\pi R} D_x \sum_{n=0}^{\infty} \sum_{m=-n}^n I_{n,m}^*(x - x_0) O_{n,m}(hy^{\text{KT}} - x_0) \right) \Bigg|_{x=x_i} \\ &= \frac{|y^{\text{KT}}|}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n F_{x_0}^{\wedge, O}(n, m) O_{n,m}(hy^{\text{KT}} - x_0) \end{aligned} \quad (22)$$

where the multipole coefficients $F_{x_0}^{\wedge, O}(n, m)$ of the cube X are given by

$$F_{x_0}^{\wedge, O}(n, m) = \sum_{\substack{i=1 \\ x_i \in X}}^N a_i \left(\frac{1}{2\pi R} D_x I_{n,m}^*(x - x_0) \right) \Bigg|_{x=x_i}. \quad (23)$$

These coefficients can be translated to other cubes via relations (19), (20) as well as (21) as long as the distance requirements are fulfilled by the construction of the decomposition of the domain into nested cubes. This first step is called *point to multipole (P2M) step*. Obviously, the infinite sum in (22) has to be truncated at degree p . This parameter essentially determines the accuracy of the algorithm.

At the end of the fast multipole cycle, i.e., after many M2M-, M2L-, L2L-translations, each cube Y possesses an inner harmonics expansion centered around the center of the cube. This expansion has to be evaluated at the targets contained by Y , i.e., the *local to targets (L2T) step* is performed:

$$\mathcal{L}_j F = F(y_j) = \left(\frac{|y^{\text{KT}}|}{R} \sum_{n=0}^p \sum_{m=-n}^n F_{x_0}^{\wedge, l}(n, m) I_{n, m}(h y^{\text{KT}} - x_0) \right) \Big|_{y=y_j}, \quad (24)$$

where the variable y is hidden by $y^{\text{KT}} = \frac{R^2}{|y|^2} y$.

In Algorithm 1 we briefly recapitulate the fast multipole algorithm (see, e.g., Carrier et al. 1988; Cheng et al. 1999 or Gutting 2007 for our specific implementation).

For the computation of the spline coefficients of the smoothing splines of Sect. 2.2, we have to consider the system of linear equations (11) instead of (8).

This means that we have to add $\beta \sum_{i=1}^N a_i (\mathbf{C}^{-1})_{ij}$ to the matrix-vector product that is computed by the FMM. In order to keep a fast algorithm, the matrix \mathbf{C}^{-1} has to allow a fast summation method or \mathbf{C} has to be a sparse matrix. The trivial cases where \mathbf{C} is a diagonal matrix can also be included in the direct evaluation step of the fast multipole algorithm.

3.5 Acceleration of the Translations

Following the ideas of White and Head-Gordon (1996) (see also Greengard and Rokhlin 1997; Cheng et al. 1999), the multipole-to-multipole (M2M) and the local-to-local (L2L) translations can both be accelerated by using Wigner rotation matrices (cf., e.g., Edmonds 1964; Biedenharn and Louck 1981; Varšalovič et al. 1988; Choi et al. 1999). With these rotations, the shift direction becomes the ε^3 -axis; we shift there and rotate back. This reduces the numerical costs from $\mathcal{O}(p^4)$ in the M2M and L2L steps to $\mathcal{O}(p^3)$, since each rotation requires an effort of $\mathcal{O}(p^3)$ and the shift along the ε^3 -axis is given by

$$\tilde{F}_{x_1}^{\wedge, 0}(n', m') = \sum_{n=|m'|}^{n'} \tilde{F}_{x_0}^{\wedge, 0}(n, m') \frac{|x_0 - x_1|^{n'-n}}{(n' - n)!}, \quad m' = -n', \dots, n',$$

Algorithm 1 Fast Multipole Algorithm

Input:

- A set of points $x_i \in \overline{\Sigma}^{\text{ext}}$ (often $x_i \in \Sigma$), $i = 1, \dots, N$,
- a set of coefficients a_i , $i = 1, \dots, N$,
- the choice of the type of the reproducing kernel $K_{\mathcal{H}}$ (singularity or Abel-Poisson with the parameter h and the radius of the Runge sphere R),
- a set of evaluation points $y_j \in \overline{\Sigma}^{\text{ext}}$, $j = 1, \dots, M$,
- the degree of the multipole expansion p ,
- the maximal number of points per cube m .

Aim: compute the sum

$$F(y_j) = \sum_{i=1}^N a_i K_{\mathcal{H}}(x_i, y_j) \quad \text{for each } j = 1, \dots, M.$$

Initialization:

- Compute the targets $hy_j^{\text{KT}} = h \frac{R^2}{|y_j|^2} y_j$, $j = 1, \dots, M$.
- Create a bounding box that contains all points and all targets, build the adaptive octree, and sort in all points and targets (as described in the beginning of Sect. 3.2). Set L as the maximum level, and eliminate all empty cubes.
- Determine list 1 to list 4 of Definition 7. Create a list of all cubes of level l for each level $l = 0, \dots, L$. Collect all leaves in a list.
- Allocate memory for the different expansion coefficients of each cube X : multipole expansion (coefficient vector M_X), local expansion (coefficient vector L_X).

Fast Multipole cycle:

1. Generation of the multipole coefficients:

For all leaves X : P2M, i.e., compute the multipole coefficients M_X of the multipole expansion up to degree p around the center of X from the points in X as in (23).

For level $l = L - 1, \dots, 2$: M2M, i.e., translate the multipole coefficients of the children of X to X itself for all cubes X of level l via (19).

2. Interaction phase for list 4:

For level $l = 2, \dots, L$: for all cubes X of level l : compute the expansion coefficients of an inner harmonics expansion around the center of X from the points in Y for all cubes Y of list 4 of X and add them to L_X – or use direct evaluation of the kernel sum corresponding to the points in Y to obtain the result at the targets in X if the number of targets in $X \leq p^2$ and X is a leaf.

3. Multipole-to-local translation:

For level $l = 2, \dots, L$: for all cubes X of level l : use (20) to translate M_X to L_Y for all cubes Y in list 2 of X .

4. Translation of the inner harmonics expansions:

For level $l = 2, \dots, L - 1$: L2L, i.e., translate the local coefficients L_X to the children of X (if there are any) via (21) and add the resulting coefficients to L_Z where Z denotes the corresponding child of X for all cubes X of level l .

5. Evaluation of the expansions and direct interaction:

For all leaves X : L2T, i.e., evaluate the inner harmonics expansion of X at all targets in X as in (24). Store the result in F .

(continued)

For all cubes Y in list 1 of X : P2T, i.e., add the direct evaluation of the kernel sum corresponding to the points in Y at the targets in X to F .

For all cubes Y in list 3 of X : evaluate the multipole expansion around the center of Y (coefficients M_Y) at the targets in X and add the results to F – or use direct evaluation of the kernel sum corresponding to the points in Y to add the result at the targets in X to F if the number of points in $Y \leq p^2$ and Y is a leaf.

6. Reverse the effects of the Kelvin transformation:

$$\tilde{F}_j = \frac{|y_j^{KT}|}{R} F_j \text{ for } j = 1, \dots, M.$$

Return the result \tilde{F} .

for M2M (the tilde indicates that we are dealing with rotated coefficients) and by

$$\tilde{F}_{x_3}^{\wedge, I}(n', m') = \sum_{n=n'}^p \tilde{F}_{x_2}^{\wedge, I}(n, m') \frac{|x_3 - x_2|^{n-n'}}{(n - n')!}, \quad m' = -n', \dots, n',$$

for L2L requiring also an effort of $\mathcal{O}(p^3)$. For a detailed description, we refer to White and Head-Gordon (1996) or Gutting (2007) with all technical details.

For the M2L translation, we follow the idea of Greengard and Rokhlin (1997) and Cheng et al. (1999) by replacing it with *exponential translations* which are based on the representation

$$\begin{aligned} \frac{1}{|x - y|} &= \frac{1}{2\pi} \int_0^\infty e^{-\lambda(x_3 - y_3)} \int_0^{2\pi} e^{i\lambda((x_1 - y_1) \cos \alpha + (x_2 - y_2) \sin \alpha)} d\alpha d\lambda \\ &= \sum_{k=1}^{s(\varepsilon)} \frac{w_k}{M_k} \sum_{j=1}^{M_k} e^{-\lambda_k(x_3 - y_3)} e^{i\lambda_k((x_1 - y_1) \cos \alpha_{j,k} + (x_2 - y_2) \sin \alpha_{j,k})} + \mathcal{O}(\varepsilon) \end{aligned} \tag{25}$$

for points x, y whose Cartesian coordinates satisfy $1 \leq x_3 - y_3 \leq 4$ as well as $0 \leq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \leq 4\sqrt{2}$. The inner integral is discretized using the trapezoidal rule, i.e., $\alpha_{j,k} = \frac{2\pi j}{M_k}$, $j = 1, \dots, M_k$, and the outer integral is treated with the integration weights w_k and the integration points λ_k , $k = 1, \dots, s(\varepsilon)$ for a chosen accuracy ε . w_k, λ_k, M_k can be found in Greengard and Rokhlin (1997), Yarvin and Rokhlin (1998), and Cheng et al. (1999). The total number of numerical

integration points, i.e., the number of exponential functions, is $S(\varepsilon) = \sum_{k=1}^{s(\varepsilon)} M_k \in$

$\mathcal{O}(p^2)$, whereas $s(\varepsilon) \in \mathcal{O}(p)$ and takes the role of p in determining the accuracy of the integration.

Since outer harmonics are related to the single pole by Hobson’s formula (cf. Hobson 1965), a multipole expansion of F with coefficients $F_c^{\wedge, O}(n, m)$, center c , and accuracy ε can be transformed by (25) into a series of exponentials (multipole-to-exponential step, briefly M2X)

$$F(x) = \sum_{k=1}^{s(\varepsilon)} \sum_{j=1}^{M_k} W(k, j) e^{-\lambda_k \frac{x_3 - c_3}{d}} e^{i \lambda_k \left(\frac{x_1 - c_1}{d} \cos \alpha_{j,k} + \frac{x_2 - c_2}{d} \sin \alpha_{j,k} \right)} + \mathcal{O}(\varepsilon), \tag{26}$$

with the coefficients given by

$$W(k, j) = \sum_{n=0}^p \sum_{m=-n}^n \frac{F_c^{\wedge, O}(n, m)}{d^{n+1}} \frac{w_k}{M_k} \lambda_k^n i^m e^{i m \alpha_{j,k}} \tag{27}$$

for $k = 1, \dots, s(\varepsilon)$, $j = 1, \dots, M_k$. It is required that c is the center of a box of edge length d containing the points x and the series of exponentials is valid for points y with $d \leq x_3 - y_3 \leq 4d$ and $0 \leq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \leq 4\sqrt{2}d$. Such exponential expansions can be shifted very efficiently to the new center b , i.e., to

$$F(x) = \sum_{k=1}^{s(\varepsilon)} \sum_{j=1}^{M_k} V(k, j) e^{-\lambda_k \frac{x_3 - b_3}{d}} e^{i \lambda_k \left(\frac{x_1 - b_1}{d} \cos \alpha_{j,k} + \frac{x_2 - b_2}{d} \sin \alpha_{j,k} \right)}$$

with the new coefficients

$$V(k, j) = W(k, j) e^{-\lambda_k \frac{b_3 - c_3}{d}} e^{i \lambda_k \left(\frac{b_1 - c_1}{d} \cos \alpha_{j,k} + \frac{b_2 - c_2}{d} \sin \alpha_{j,k} \right)} \tag{28}$$

for $k = 1, \dots, s(\varepsilon)$, $j = 1, \dots, M_k$. This exponential to exponential shift is abbreviated by X2X.

Afterwards, the exponential expansion is transformed back into an inner harmonics expansion completing the M2L translation step:

$$F(x) = \sum_{n=0}^p \sum_{m=-n}^n F_c^{\wedge, I}(n, m) I_{n,m}(x - c) + \mathcal{O}(\varepsilon), \tag{29}$$

where the coefficients can be computed by this so-called exponential to local (X2L) step

$$F_c^{\wedge, I}(n, m) = \sum_{k=1}^{s(\varepsilon)} \sum_{j=1}^{M_k} W(k, j) \frac{(-\lambda_k)^n}{d^n} i^m e^{-i m \alpha_{j,k}} \tag{30}$$

for $n = 0, \dots, p$, $m = -n, \dots, n$. Obviously, the same geometric restrictions hold as before.

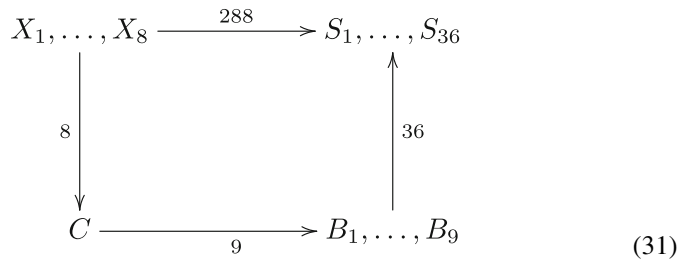
The restrictions on the positions of x and y mean that the exponential translations are applicable for cubes in list 2 that are situated above the current cube with another cube in between. However, by combining the idea with rotations of the multipole expansion using again the Wigner rotation matrices, the exponential translation can

substitute the M2L translation for all cubes in list 2. Therefore, the list of all well-separated cubes (list 2) is split into six directional lists (up, down, North, South, East, and West), and instead of M2L, the following sequence of transformations is used: (rotation), M2X (27), X2X (28), and X2L (30), (inverse rotation).

Each exponential shift requires numerical costs of $\mathcal{O}(S(\varepsilon)) = \mathcal{O}(p^2)$, and the rotations can be applied using $\mathcal{O}(p^3)$ operations (as do the M2X and X2L steps). Thus, this improves the performance compared to the M2L step's $\mathcal{O}(p^4)$ effort. Each cube needs to allocate some additional memory for an outgoing and an incoming set of coefficients of the sum of exponentials.

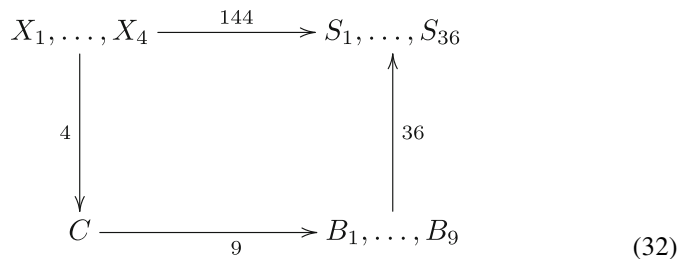
Moreover, we can save translations by recombination (see Greengard and Rokhlin 1997; Cheng et al. 1999; Gutting 2007). If several cubes have the same target cubes, intermediate stops are performed. No additional memory is required since we can use the available storage of cubes at a lower level.

In 3D we perform 8 translations from X_1, \dots, X_8 to their common parent cube C , 9 translations from C to B_1, \dots, B_9 which are the parent cubes of S_1, \dots, S_{36} instead of 288 translations from X_1, \dots, X_8 to S_1, \dots, S_{36} as summarized in the following diagram:



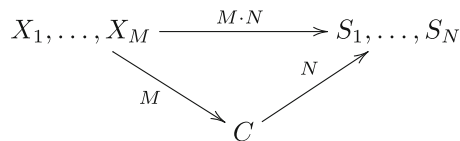
Not only expansions of X_j , but also of other cubes that have the children of B_k as target cubes, are collected in the parent cubes B_k and are shifted to the cubes S_j after all contributions are added. Thus, we only need to perform the translations from B_k to its children once. This further reduction is the reason that we use two intermediate stations for the translation. We call this the first stage of recombination.

Similar considerations can be used in other situations where fewer cubes have common target cubes (second stage of recombination). The corresponding diagram for the 3D case is given by



As before, the final translation is executed only after all contributions arrived in the cubes B_k . These two stages already cover the complete list for the directions up and down. For the other four lists, we introduce a third stage. Obviously, we only consider cubes that have not yet been treated by the first two stages.

If a cube X has a cube Y in one of its directional lists, we collect all other children of X 's parent C which also have Y in the corresponding directional list in the combination list (altogether M cubes). If $M > 1$, we collect the maximal number of common cubes in the corresponding directional lists of all cubes of the combination list in the target cube list (altogether N cubes). If $N > 1$, all expansions from cubes in the combination list are shifted to their parent cube C and then to the cubes of the target cube list. The number of translations is reduced from $N \cdot M$ to $N + M$.



Finally, any remaining cubes are treated individually as without the recombination technique. This procedure can significantly reduce the number of necessary translations. The algorithm is described and analyzed in full detail in Gutting (2007).

It should also be noted that there are several symmetries in the coefficients of the exponential expansion since we are dealing with a real-valued function F in (26). They are used to further reduce the numerical costs (cf. Greengard and Rokhlin 1997; Cheng et al. 1999).

4 Numerical Tests and Results

At first, the effect of the recombination described in the end of Sect. 3.5 is investigated for a fully populated bounding box. Obviously, this is the ideal case maximizing the effect of the recombination. The increased efficiency of the X2X step for the various stages of recombination (see Sect. 3.5; by stage 0 we indicate the case without recombination) is documented by Tables 1 and 2. Note that the accuracy $s(\varepsilon) = 26$ and the corresponding values for M_k , $S(\varepsilon)$ from Cheng et al. (1999) are chosen for this test. The well known symmetries in the exponential coefficients (see Gutting 2007 for a detailed derivation) are also used to reduce the computational costs. For a rather sphere-like setting as in the discrete version of the Dirichlet boundary value problem (Problem 3) or in an interpolation/approximation problem on the Earth's surface, the savings reduce to about 30% because of the adaptive construction of the octree. This can be seen in Table 3 where an adaptive decomposition is computed for points randomly distributed between two spheres of radius 6,356 and 6,378 km. The targets are the Kelvin transforms of these points

Table 1 Number of exponential translations depending on the stages in the recombination technique for the fully populated octree with levels 2–5. In brackets in the first column, the number of cubes in this level is given; the average number of translations per cubes is listed in brackets in the other columns

Level	Stage 0	Stage 1	Stage 2	Stage 3
2 (64)	3,096(48.4)	1,648(25.8)	1,000(15.6)	902(14.1)
3 (512)	53,352(104.2)	27,792(54.3)	16,056(31.4)	13,602(26.6)
4 (4,096)	584,136(142.6)	302,176(73.8)	171,640(41.9)	141,050(34.4)
5 (32,768)	5,398,920(164.8)	2,784,960(85.0)	1,570,680(47.9)	1,271,850(38.8)

Table 2 Computing times (in seconds) corresponding to the exponential translations of Table 1

Level	Stage 0	Stage 1	Stage 2	Stage 3
2	0.61	0.32	0.20	0.18
3	9.96	5.27	3.11	2.67
4	107.89	56.55	32.87	27.34
5	1,040.81	548.03	330.88	263.01

Table 3 Computing times (in seconds) for the exponential translation without and with recombination (all three stages). Note that the listed times include the M2X and X2L transformations in both cases

Maximal level	Cubes	Leaves	Without recombination	With recombination
4	1,313	1,112	30.8	24.21
4	1,545	1,280	37.37	29.24
5	1,682	1,384	41.62	32.26
5	2,630	2,200	58.74	47.81

with respect to a sphere of radius $R = 6,352$ km and the parameter $h = 0.95$. The adaptive FMM used the parameters $p = 23$, $s(\varepsilon) = 26$, $S(\varepsilon) = 670$, $m = 130$.

Now the truncation degree p is investigated for different accuracies of the exponential translation $s(\varepsilon)$. We increase p while $s(\varepsilon)$ is kept fixed and determine when the integration error of the exponential translation dominates the truncation error. This leads to the choices of p for different levels of $s(\varepsilon)$ given by Table 4.

For the remaining tests in this section, we consider the following test scenario: We take latitudes and longitudes of the points of a so-called spiral grid (cf. Rakhmanov et al. 1994) and combine these with a random radius between 6,356.7 and 6,378.1 km. As radius of the Runge sphere, $R = 6,356$ km is chosen and $h = 0.95$. The spline coefficients a_i are just random numbers between -1 and 1 . We compute

$$S(x_j) = \sum_{i=1}^N a_i \mathcal{L}_j \mathcal{L}_i K_{\mathcal{H}}(\cdot, \cdot) = \sum_{i=1}^N a_i K_{\mathcal{H}}(x_i, x_j), \quad j = 1, \dots, N,$$

Table 4 Resulting truncation degrees p for different $s(\varepsilon)$ for the two types of kernels

$s(\varepsilon)$	Singularity kernel	Abel-Poisson kernel
8	4	5
17	12	13
26	23	25

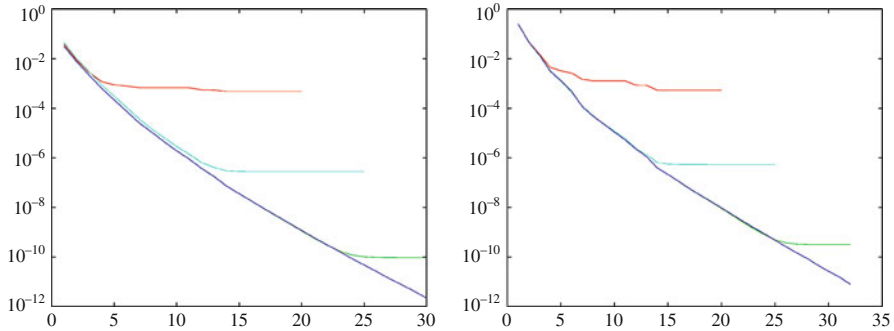


Fig. 3 Error induced by the FMM for the singularity kernel (left) and Abel-Poisson kernel (right) with exponential translations of different accuracies for $s(\varepsilon) = 8$ (red line), $s(\varepsilon) = 17$ (cyan line), $s(\varepsilon) = 26$ (green line). For comparison, we include the standard M2L translation (blue line) with increasing truncation degree p as abscissae

Table 5 Chosen maximal numbers of points m per cube for the singularity kernel and the Abel-Poisson kernel and the different error levels

$s(\varepsilon)$	Singularity kernel	Abel-Poisson kernel
8	85	75
17	130	140
26	380	240

with exponential translations of different accuracies and without them (using standard M2L translations). We compare the results to the direct summation without FMM. The error behavior can be seen in Fig. 3, and together with many further tests (cf. Gutting 2007), we obtain the values of Table 4.

Finally, the maximal number of points or targets per cube m has a strong influence on the adaptive octree construction and the performance of the FMM. If m is too small, there are many cubes each containing only very few points. Thus, the kernel expansion coefficients no longer combine the information of enough points to be efficient. If m is too large, there are only few cubes each with a large number of points. This means that far too often instead of kernel expansion direct interaction is used. After many empirical tests (cf. Gutting 2007), we came to the conclusion that the choices for m given by Table 5 lead to a good performance in our implementation.

After these optimizations of the parameters of the fast multipole algorithm, we can compare its performance with direct computation and find the break-even points of our implementation, i.e., the minimal number of points that is necessary for our algorithm to be faster than the direct approach (see Table 6). Note that such results are always very dependent on the implementation.

Table 6 Break-even points for the singularity kernel and the Abel-Poisson kernel

$s(\varepsilon)$	Singularity kernel	Abel-Poisson kernel
8	530	360
17	1,160	960
26	2,670	2,250

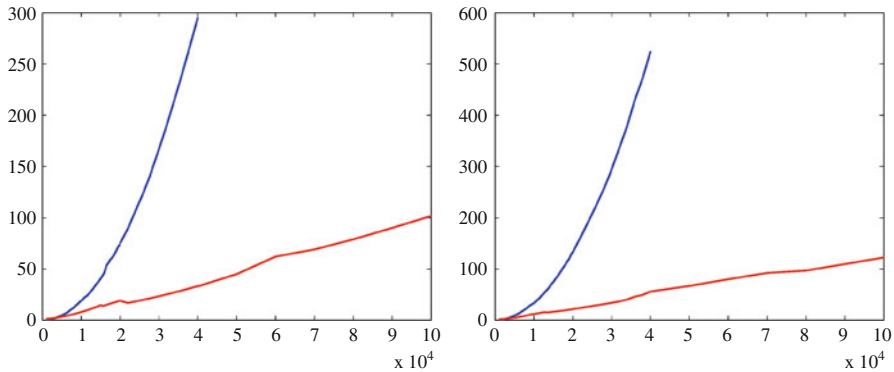


Fig. 4 Break-even points by comparison of computation times (in seconds) for direct (blue line) and FMM-accelerated (red line) computation (left: singularity kernel, right: Abel-Poisson kernel), the number of points forms the abscissae

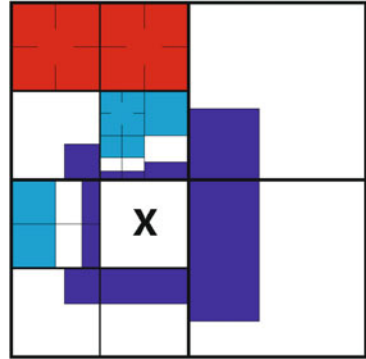
Furthermore, the linear asymptotic behavior which we expect from the FMM becomes obvious in Fig. 4 compared to the quadratic behavior of the direct approach. Our implementation turns out to be efficient even for rather small problem sizes. In general, the Abel-Poisson kernel requires some more computation time since it leads to a more difficult P2M step.

5 Overlapping Domain Decomposition Preconditioner

The fast multipole algorithm already provides us with an adaptive decomposition of the computational domain which we also use to improve the condition of the system of linear equations (8) by using a Schwarz algorithm as preconditioner (see Chan and Mathew 1994; Smith et al. 1996 and the references therein for an overview). The basic idea of the Schwarz algorithms is to split up the computational domain into several (often overlapping) parts, solve the resulting smaller problems for each part, and combine the partial solutions. We intend to use only one decomposition for both the FMM and the Schwarz algorithm.

There are two main variants that have different ways to update the residual: In the multiplicative variant, the residual is updated after each single solution of a subproblem. In the additive variant, the solutions of the subproblems are merged after all of them are computed, and the merged result updates the residual (cf. Beatson et al. 2000; Zhou et al. 2003 for use of these methods in radial basis function interpolation and Hesse 2002 for use of the multiplicative variant in the context

Fig. 5 Two-dimensional illustration of the overlap (dark blue area) for an adaptive decomposition of the domain. The red and light blue areas belong to list 2 and 3, respectively



of harmonic splines). The multiplicative variant possesses the advantage of not requiring the full matrix for the update of the residuals, but this is less relevant in the presence of the FMM which provides a very fast way to (approximately) multiply with the full matrix. Moreover, the total number of update steps is higher for the multiplicative variant. Thus, we choose the additive Schwarz algorithm together with a coarse grid correction and use a residual update of Hybrid-II type (by the categorization of Smith et al. 1996).

Since all the decomposing of the domain and sorting of points is already part of the FMM, the only additional work in the initialization of the preconditioner is the determination of points in the overlapping parts of the subdomains which are exactly given by the leaves of the octree. The amount of the overlap is controlled by the parameter ϑ_{ov} , and the area is a part of the directly adjacent cubes, i.e., cubes in list 1 as set in Definition 7 (white in Fig. 5). The overlap is the blue area in Fig. 5. Note that $\vartheta_{ov} = 0$, i.e., no overlap at all, is a valid choice. In our problems we usually do not assume a structured point distribution. The idea for establishing a coarse grid from our scattered data points is to choose from each subproblem the point that is closest to the center of the domain and add that point to the coarse grid. Note that for the preconditioner, we only consider points and ignore the so-called targets of Sect. 3.2. If a cube only contains targets and none of the points x_1, \dots, x_N , neither overlap nor coarse grid points are determined for it. After these initialization steps that need to be computed only once, we can investigate the preconditioning cycle which yields $v = \mathbf{M}^{-1}b$ for a given residual b .

At first, we introduce the *restriction matrix* for each subdomain. Let $M \in \mathbb{N}$ be the total number of subdomains and the subdomain X_r , $r = 1, \dots, M$, contain N_r points in the cube and its corresponding overlap. By σ_r , we denote the permutation of the numbers $1, \dots, N$ that yields the points $x_{\sigma_r(i)}$ of X_r for the indices $i = 1, \dots, N_r \leq N$. For $i = N_r + 1, \dots, N$, the permutation σ_r is supposed to give the remaining indices between 1 and N such that $\sigma_r(i) < \sigma_r(i + 1)$ for all $i = N_r + 1, \dots, N - 1$. With the help of the canonical basis vectors in \mathbb{R}^N and the index permutation σ_r , the restriction matrix \mathbf{R}_r is defined as

$$\mathbf{R}_r = \begin{pmatrix} e_{\sigma_r(1)}^T \\ \vdots \\ e_{\sigma_r(N_r)}^T \end{pmatrix} \in \mathbb{R}^{N_r \times N}. \quad (33)$$

The restriction matrices reduce a vector $x \in \mathbb{R}^N$ to a vector in \mathbb{R}^{N_r} consisting only of components of x related to X_r . The corresponding transposed matrix $\mathbf{R}_r^T \in \mathbb{R}^{N \times N_r}$ is called the *prolongation matrix* in the multiplicative variant. The restriction/prolongation matrices $\mathbf{R}_0, \mathbf{R}_0^T$ for the coarse grid are defined analogously using the corresponding set of points. In the additive variant with overlapping subdomains, $\mathbf{R}_r^T, r = 1, \dots, M$, stands for the operation of fitting together solutions corresponding to the subdomains. Several possibilities are available (see Gutting 2007 and the references therein), and we choose to only update components that correspond to points of the cube in X_r without the overlap. The overlapping part still plays a role in solving the subproblems, i.e., in \mathbf{R}_r and \mathbf{A}_r . Obviously, these restriction and prolongation operators are only written as matrices for the description of the algorithm; their effect is implemented in the form of index operations.

Algorithm 2 Overlapping Additive Schwarz Preconditioner

Initialization step:

Find points $x_i, i = 1, \dots, N_r$, in each domain including overlap and set matrices

$$\mathbf{A}_r = (K_{\mathcal{H}}(x_i, x_j))_{i,j=1,\dots,N_r}, \quad r = 1, \dots, M. \quad (34)$$

Find the coarse grid and build the corresponding matrix $\mathbf{A}_0 = (K_{\mathcal{H}}(x_i, x_j))$, where $i, j = 1, \dots, N_0$ correspond to the indices of the N_0 points of the coarse grid. The residual b is given.

Preconditioning cycle:

For $r = 1, \dots, M$: Solve $\mathbf{A}_r z_r = \mathbf{R}_r b$ to obtain z_r .

Update at the inner points: $v = \sum_{r=1}^M \mathbf{R}_r^T z_r$.

Update the coarse residual and perform coarse grid correction:

$$v_{\text{final}} = v + \mathbf{R}_0^T \mathbf{A}_0^{-1} \mathbf{R}_0 (b - \mathbf{A}v), \quad (35)$$

where \mathbf{A} denotes the full matrix corresponding to all N points.

Remark 11. The computation of $\mathbf{R}_0 \mathbf{A} v$ in (35) stands for the product of an $N_0 \times N$ -matrix with a vector in \mathbb{R}^N . Since $N_0 \ll N$, this matrix-vector product is computed directly in our implementation even though it is of the type for which the FMM works as fast summation method.

Our implementation solves the subproblems directly, i.e., the matrices \mathbf{A}_r are factorized once in the initialization step.

Table 7 Splines using the singularity and the Abel-Poisson kernel with $N = 8,800$ or $N = 11,200$ points. Computation times (in seconds) and the number of iterations in brackets. The first line uses a direct solver for comparison; the second line uses GMRES without preconditioning. Note that we restrict the GMRES algorithm to a maximum of 300 iterations. If no convergence is achieved by then, we write >300

ϑ_{ov}	Singularity (8,800 points)	Singularity (11,200 points)	Abel-Poisson (8,800 points)	Abel-Poisson (11,200 points)
Direct	331.32	580.30	291.81	577.53
No PC.	1,020.7 (>300)	1,187.84 (>300)	522.76 (84)	747.14 (101)
0.1	108.20 (29)	149.50 (35)	90.29 (13)	131.38 (16)
0.2	52.16 (13)	84.45 (19)	52.48 (7)	63.06 (7)
0.3	38.00 (9)	52.68 (11)	40.26 (5)	50.27 (5)
0.4	32.01 (7)	37.82 (7)	34.71 (4)	50.63 (5)
0.5	26.20 (5)	37.62 (6)	36.23 (4)	46.41 (4)
0.6	29.20 (5)	39.91 (5)	39.33 (4)	53.16 (4)

Algorithm 2 is applied as preconditioner in the well known preconditioned GMRES algorithm for the iterative solution of systems of linear equations (cf. Saad and Schultz 1986 for the first formulation of GMRES or, e.g., Saad 2003). The matrix-vector products with the full matrix \mathbf{A} that occur in GMRES are accelerated with the fast multipole algorithm of Sect. 3.

The overlap parameter ϑ_{ov} which controls the size of the overlapping parts of the domains (see Fig. 5) plays an important role in the effectiveness of the preconditioner. Note that the number of overlap points is not directly controlled this way. On the one hand, a larger overlapping part, i.e., more overlap points, generally reduces the number of iterations, because the efficiency of the additive Schwarz preconditioner rises. On the other hand, the number of points related to each domain also grows, i.e., the effort for each domain increases, and hence the iterations take more time.

We investigate the following test scenario to determine a good balance. The Earth's surface is described by the TerrainBase model (see Hastings and Row 1997), and we choose the latitudes and longitudes of the data points from a so-called spiral grid (cf. Rakhmanov et al. 1994), i.e., the points are approximately equidistributed on the surface for this test. As gravitational data to interpolate, we use the EGM96 model (cf. Lemoine et al. 1998). The fast multipole algorithm is run with $s(\varepsilon) = 17$ and corresponding values for p and m (see Sect. 4). The spline parameters are $R = 6,352$ km and $h = 0.92$ for both kernels. From Table 7, we find that the differences for $\vartheta_{ov} \geq 0.4$ are rather small. Moreover, the disadvantages of a too large overlap become more and more evident for more points, in particular more points in the cubes, e.g., resulting from $s(\varepsilon) = 26$ (see Table 5). Therefore, we use the preconditioner with $\vartheta_{ov} = 0.4$ in our calculations in Sect. 6. Further parameter studies and tests of other update variants can be found in Gutting (2007).

For smoothing splines, similar restrictions on the matrix \mathbf{C} in (11) hold as for the FMM (see the end of Sect. 3.4). However, the system of linear equations (11) is typically much better conditioned than (8) because of the smoothing.

6 Numerical Tests and Results

At first, we consider a global test example putting together all previous parts. In this section, the FMM is used with $s(\varepsilon) = 26$ and the corresponding parameters (see Sect. 4), and the preconditioner is set up with the overlap parameter $\vartheta_{\text{ov}} = 0.4$ as explained in Sect. 5.

To determine random points on the Earth's surface, we compute a uniform distribution on a sphere. Since the Earth's surface is close to a sphere, the distribution can be expected to be still close to uniform. Of course, even for uniformly distributed random points, data gaps can occur leading to complications. As test data, we simply evaluate the EGM96 (using degrees 3–100 here) at $N = 100,000$ random points on the Earth's surface. In Fig. 6, the spline of the gravitational potential and distribution of the absolute error of the spline interpolation can be seen. Figure 7 provides a detailed view of the errors at some data gaps.

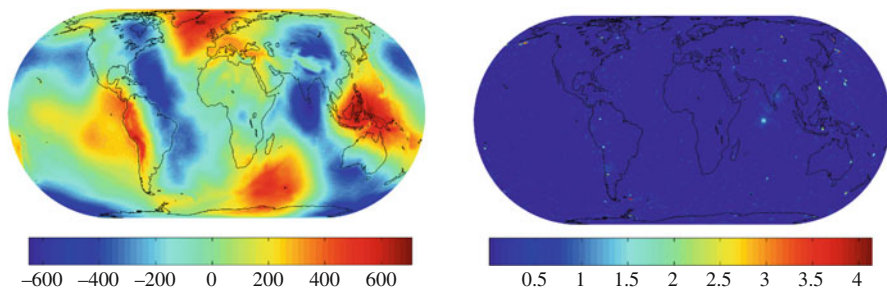


Fig. 6 Spline interpolation of the gravitational potential (*left*) and absolute error distribution (*right*), both in m^2/s^2

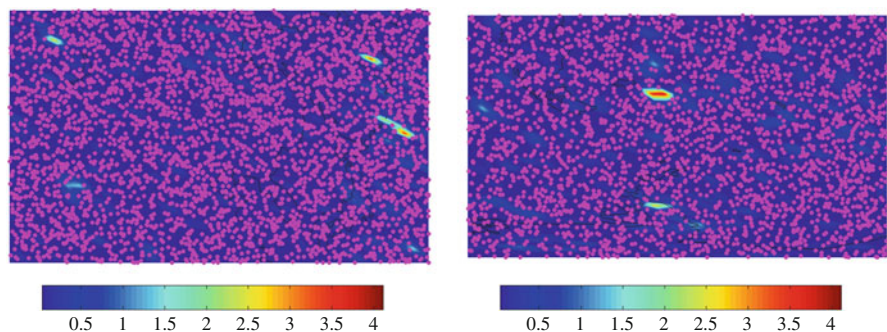


Fig. 7 A closer look at some data gaps where the largest error (in m^2/s^2) occurs in Fig. 6 (*right*). The data points are added in magenta

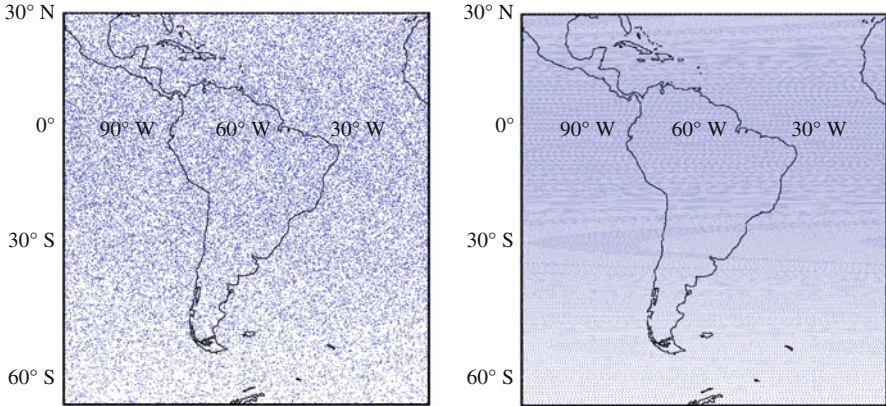


Fig. 8 Local examples: $N = 48,749$ randomly distributed points (*left*) and $N = 48,744$ points based on a spiral grid (*right*) in the region of the Earth's surface

The singularity kernel is computed with $h = 0.98$ and $R = 6,352$ km in just 18 iterations and leads to a mean absolute error of $\varepsilon_{\text{mabs}} = 0.0413 \text{ m}^2/\text{s}^2$ and a maximal absolute error of $\varepsilon_{\text{max}} = 4.1442 \text{ m}^2/\text{s}^2$.

For our further considerations, we no longer work on the whole surface of the Earth but restrict ourselves to a region containing South America (65°S to 30°N , 110°W to 10°W), and only the points within this region are interpolated. As test data, the EGM96 (using degrees 16–200) is evaluated at $N = 48,749$ almost uniformly distributed random points (constructed as before) in the region on the Earth's surface (see Fig. 8 (*left*)).

The resulting spline with the singularity kernel with $h = 0.9875$ and $R = 63,544$ km required 32 iterations to obtain the coefficients, and its mean absolute error is $\varepsilon_{\text{mabs}} = 0.0353 \text{ m}^2/\text{s}^2$. The maximal error $\varepsilon_{\text{max}} = 8.0492 \text{ m}^2/\text{s}^2$ and occurs at the boundary of the region (see Fig. 9). Thus, we ignore 2.5 % of the region at each boundary (see Fig. 10) which reduces both mean and maximal errors to $\tilde{\varepsilon}_{\text{mabs}} = 0.0280 \text{ m}^2/\text{s}^2$ and $\tilde{\varepsilon}_{\text{max}} = 3.6544 \text{ m}^2/\text{s}^2$, respectively (compare also Figs. 10 to 9). It should be noted that the error naturally is largest at the boundary since there are no more data points beyond it. Therefore, one can consider everything outside the region under consideration as one huge data gap.

Finally, we consider another regional example with the Abel-Poisson kernel (with $h = 0.98$ and $R = 6,354$ km) using the spiral grid of Rakhmanov et al. (1994) mapped to the Earth's surface as before (see the test scenarios in Sect. 5). The points ($N = 48,744$) are illustrated in Fig. 8 (*right*), and as data we evaluate again the EGM96 (using degrees 16–360). Our algorithm requires 15 iterations, and the errors are $\varepsilon_{\text{mabs}} = 0.0199 \text{ m}^2/\text{s}^2$ and $\varepsilon_{\text{max}} = 10.7790 \text{ m}^2/\text{s}^2$ where the maximal error again occurs at the boundary of the region (see Fig. 11). Cutting off again 2.5 % of the region at the boundary (see Fig. 12 and compare it to Fig. 11), these errors reduce to $\tilde{\varepsilon}_{\text{mabs}} = 0.0046 \text{ m}^2/\text{s}^2$ and $\tilde{\varepsilon}_{\text{max}} = 0.9984 \text{ m}^2/\text{s}^2$. Obviously, the more regular distribution of points removes the errors that result from data gaps and leads to better results.

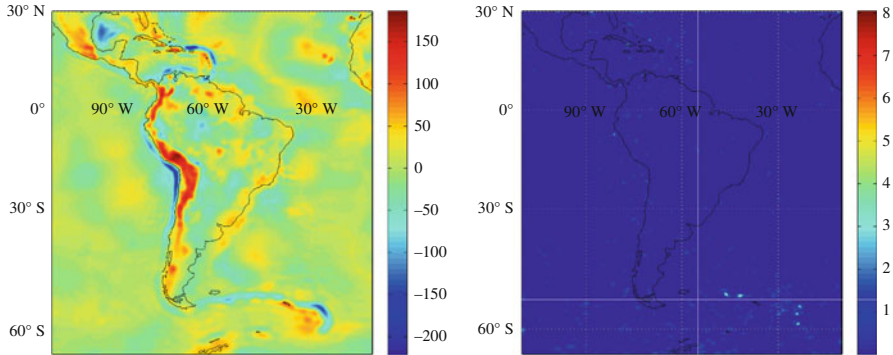


Fig. 9 Spline interpolation of the gravitational potential (*left*) and absolute error distribution (*right*) using the singularity kernel and the interpolation points in Fig. 8 (left), both in m^2/s^2

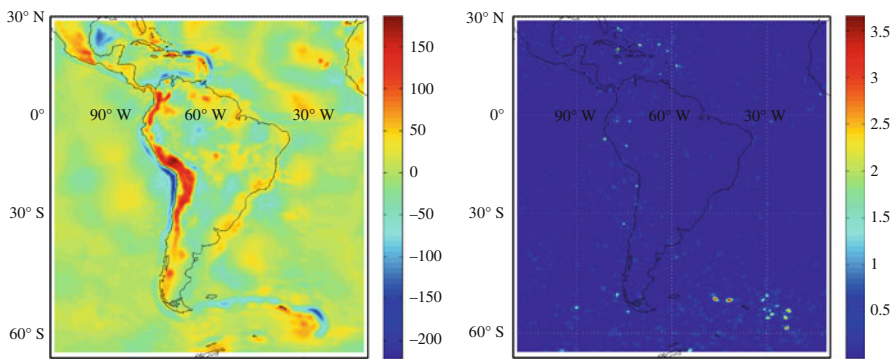


Fig. 10 Same results as in Fig. 9, but 2.5 % of the area at the boundary of the region have been removed

7 Conclusion

Combining the FMM and the additive Schwarz algorithm as preconditioner in the iterative algorithm GMRES proves to be an efficient solution strategy that can treat interpolation problems and Dirichlet boundary value problems with many data points on regular surfaces (e.g., the actual topography of the Earth). It should be pointed out that our approach is not restricted to a global treatment but also applies to regional domains as shown by the numerical examples. This can lead to a local improvement of the gravitational field in areas of particular interest. A global model (e.g., in terms of spherical harmonics) should first be subtracted from the data and combined afterwards with the spline solution. In general, the singularity kernel leads to results slightly faster than the Abel-Poisson kernel.

The spline approach naturally includes spherical boundaries as a special case and can be extended to spline smoothing with some restrictions (see Sect. 2.2 and

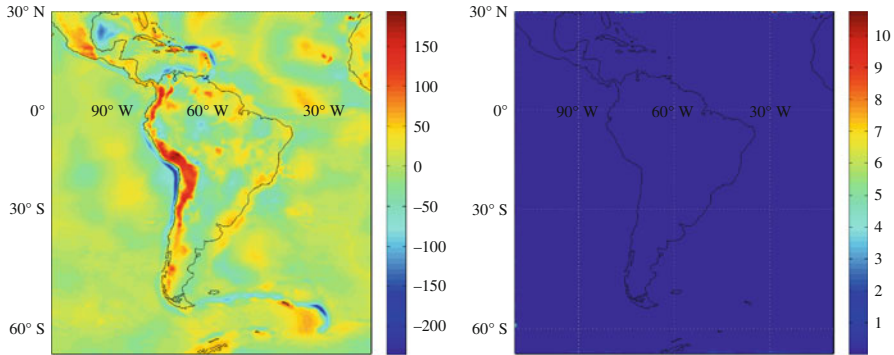


Fig. 11 Spline interpolation of the gravitational potential (*left*) and absolute error distribution (*right*) using the Abel-Poisson kernel and the interpolation points in Fig. 8 (right), both in m^2/s^2

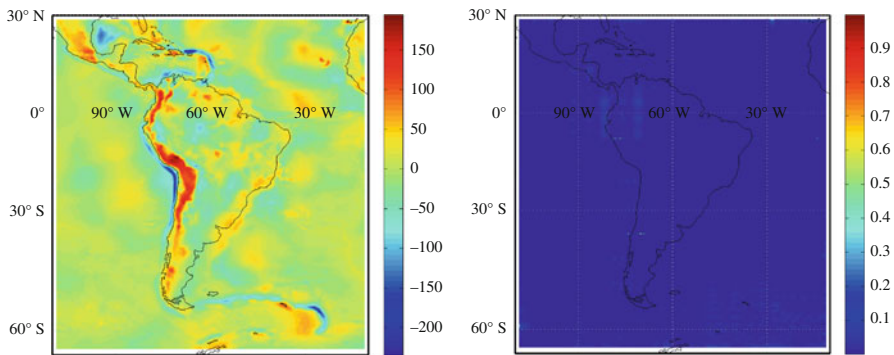


Fig. 12 Same results as in Fig. 11, but 2.5 % of the area at the boundary of the region have been removed

the end of Sect. 3.4). Note that for smoothing splines, it is well possible to leave out the preconditioner since the smoothing itself drastically improves the condition of the matrix. However, the smoothing parameter(s) plays a crucial role in this approach and must be chosen very carefully or a lot of detail information is lost to oversmoothing. The combination with parameter choice methods from ill-posed problems (cf. Bauer and Lukas 2011; Bauer et al. 2014 and the references therein) is an interesting challenge for the future.

For highly irregular distributions of data points, the spline approach reaches its limits. The largest data gap in the domain desires a small value of the parameter h , whereas the closest data points require a larger value of h to avoid ill-conditioning. Even smoothing splines cannot completely bridge this gap so far though further investigation is required. However, functional matching pursuit methods can result in better approximations (see Michel 2014a and the references therein), but so far these algorithms require large numerical costs.

References

- Aronszajn N (1950) Theory of reproducing kernels. *Trans Am Math Soc* 68:337–404
- Bauer F, Lukas MA (2011) Comparing parameter choice methods for regularization of ill-posed problems. *Math Comput Simul* 81(9):1795–1841
- Bauer F, Gutting M, Lukas MA (2014) Evaluation of parameter choice methods for regularization of ill-posed problems in geomathematics. In: Freeden W, Nashed MZ, Sonar T (eds) *Handbook of geomathematics*, 2nd edn. Springer, Heidelberg
- Beatson RK, Billings S, Light WA (2000) Fast solution of the radial basis function interpolation equations: domain decomposition methods. *SIAM J Sci Comput* 22(5):1717–1740
- Biedenharn LC, Louck JD (1981) Angular momentum in quantum physics (theory and application). *Encyclopedia of mathematics and its applications*. Addison-Wesley, Reading
- Carrier J, Greengard L, Rokhlin V (1988) A fast adaptive multipole algorithm for particle simulations. *SIAM J Sci Stat Comput* 9(4):669–686
- Chan TF, Mathew TP (1994) Domain decomposition algorithms. *Acta Numer* 3:61–143
- Cheng H, Greengard L, Rokhlin V (1999) A fast adaptive multipole algorithm in three dimensions. *J Comput Phys* 155:468–498
- Choi CH, Ivanic J, Gordon MS, Ruedenberg K (1999) Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion. *J Chem Phys* 111(19):8825–8831
- Edmonds AR (1964) *Drehimpulse in der Quantenmechanik*. Bibliographisches Institut, Mannheim
- Epton MA, Dembart B (1995) Multipole translation theory for the three-dimensional Laplace and Helmholtz equations. *SIAM J Sci Comput* 16(4):865–897
- Fengler MJ (2005) Vector spherical harmonic and vector wavelet based non-linear Galerkin schemes for solving the incompressible Navier–Stokes equation on the sphere. PhD thesis, Geomathematics Group, Department of Mathematics, University of Kaiserslautern, Shaker, Aachen
- Freeden W (1981a) On approximation by harmonic splines. *Manuser Geod* 6:193–244
- Freeden W (1981b) On spherical spline interpolation and approximation. *Math Method Appl Sci* 3:551–575
- Freeden W (1982a) Interpolation and best approximation by harmonic spline functions. *Boll Geod Sci Aff* 1:105–120
- Freeden W (1982b) On spline methods in geodetic approximation problems. *Math Method Appl Sci* 4:382–396
- Freeden W (1984a) Spherical spline interpolation: basic theory and computational aspects. *J Comput Appl Math* 11:367–375
- Freeden W (1984b) Ein Konvergenzsatz in sphärischer Spline-Interpolation. *Z f Vermessungswes.(ZfV)* 109:569–576
- Freeden W (1987a) A spline interpolation method for solving boundary value problems of potential theory from discretely given data. *Numer Methods Partial Differ Equ* 3:375–398
- Freeden W (1987b) Harmonic splines for solving boundary value problems of potential theory. In: Mason JC, Cox MG (eds) *Algorithms for approximation. The institute of mathematics and its applications, conference Series, vol 10*. Clarendon Press, Oxford, pp 507–529
- Freeden W (1999) *Multiscale modelling of spaceborne geodata*. B.G. Teubner, Stuttgart/Leipzig
- Freeden W, Gerhards C (2013) *Geomathematically oriented potential theory*. Chapman & Hall/CRC, Boca Raton
- Freeden W, Gutting M (2013) *Special functions of mathematical (geo-)physics*. Birkhäuser, Basel
- Freeden W, Michel V (2004) *Multiscale potential theory (with applications to geoscience)*. Birkhäuser, Boston/Basel/Berlin
- Freeden W, Schreiner M (2014) Special functions in mathematical geosciences: an attempt at a categorization. In: Freeden W, Nashed MZ, Sonar T (eds) *Handbook of geomathematics*, 2nd edn. Springer, Heidelberg
- Freeden W, Schreiner M, Franke R (1997) A survey on spherical spline approximation. *Surv Math Ind* 7:29–85

- Freeden W, Gervens T, Schreiner M (1998a) Constructive approximation on the sphere (with applications to geomathematics). Oxford Science Publications, Clarendon Press, Oxford
- Freeden W, Glockner O, Schreiner M (1998b) Spherical panel clustering and its numerical aspects. *J Geodesy* 72:586–599
- Glockner O (2002) On numerical aspects of gravitational field modelling from SST and SGG by harmonic splines and wavelets (with application to CHAMP data). PhD thesis, Geomathematics Group, Department of Mathematics, University of Kaiserslautern. Shaker, Aachen
- Greengard L (1988) The rapid evaluation of potential fields in particle systems. MIT, Cambridge
- Greengard L, Rokhlin V (1987) A fast algorithm for particle simulations. *J Comput Phys* 73(1):325–348
- Greengard L, Rokhlin V (1988) Rapid evaluation of potential fields in three dimensions. In: Anderson C, Greengard L (eds) *Vortex methods*. Springer, Berlin/Heidelberg/New York, pp 121–141
- Greengard L, Rokhlin V (1997) A new version of the fast multipole method for the Laplace equation in three dimensions. *Acta Numer* 6:229–269
- Gutting M (2007) Fast multipole methods for oblique derivative problems. PhD thesis, Geomathematics Group, Department of Mathematics, University of Kaiserslautern. Shaker, Aachen
- Gutting M (2012) Fast multipole accelerated solution of the oblique derivative boundary value problem. *GEM Int J Geom* 3(2):223–252
- Hastings D, Row LW III (1997) TerrainBase global Terrain model summary documentation. National Geodetic Data Center, Boulder
- Hesse K (2002) Domain decomposition methods for multiscale geopotential determination from SST and SGG. PhD thesis, Geomathematics Group, Department of Mathematics, University of Kaiserslautern. Shaker, Aachen
- Hobson EW (1965) The theory of spherical and ellipsoidal harmonics (second reprint). Chelsea Publishing Company, New York
- Keiner J, Kunis S, Potts D (2006) Fast summation of radial functions on the sphere. *Computing* 78:1–15
- Kellogg OD (1967) Foundation of potential theory. Springer, Berlin/Heidelberg/New York
- Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH, Olson TR (1998) The development of the joint NASA GSFC and NIMA geopotential model EGM96. NASA/TP-1998-206861, NASA Goddard Space Flight Center, Greenbelt
- Michel V (2013) Lectures on constructive approximation – Fourier, spline, and wavelet methods on the real line, the sphere, and the ball. Birkhäuser, Boston
- Michel V (2014a) RFMP – an iterative best basis algorithm for inverse problems in the geosciences. In: Freeden W, Nashed MZ, Sonar T (eds) *Handbook of geomathematics*, 2nd edn. Springer, Heidelberg
- Michel V (2014b) Tomography: problems and multiscale solutions. In: Freeden W, Nashed MZ, Sonar T (eds) *Handbook of geomathematics*, 2nd edn. Springer, Heidelberg
- Moritz H (2014) Classical physical geodesy. In: Freeden W, Nashed MZ, Sonar T (eds) *Handbook of geomathematics*, 2nd edn. Springer, Heidelberg
- Potts D, Steidl G (2003) Fast summation at nonequispaced knots by NFFT's. *SIAM J Sci Comput* 24(6):2013–2037
- Rakhmanov EA, Saff EB, Zhou YM (1994) Minimal discrete energy on the sphere. *Math Res Lett* 1:647–662
- Rokhlin V (1985) Rapid solution of integral equations of classical potential theory. *J Comput Phys* 60:187–207
- Saad Y (2003) Iterative methods for sparse linear systems, 2nd edn. SIAM, Philadelphia
- Saad Y, Schultz MH (1986) GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J Sci Stat Comput* 7:856–869
- Shure L, Parker RL, Backus GE (1982) Harmonic splines for geomagnetic modelling. *Phys Earth Planet Inter* 28:215–229

- Smith BF, Bjørstad PE, Gropp WD (1996) Domain decomposition (parallel multilevel methods for elliptic partial differential equations). Cambridge University Press, Cambridge
- Varšalovič DA, Moskalev AN, Chersonskij VK (1988) Quantum theory of angular momentum. World Scientific, Singapore
- Wahba G (1981) Spline interpolation and smoothing on the sphere. *SIAM J Sci Stat Comput* 2:5–16. Also errata: *SIAM J Sci Stat Comput* 3:385–386 (1981)
- Wahba G (1990) Spline models for observational data. SIAM, Philadelphia
- White CA, Head-Gordon M (1996) Rotating around the quartic angular momentum barrier in fast multipole method calculations. *J Chem Phys* 105(12):5061–5067
- Yamabe H (1950) On an extension of the Helly's theorem. *Osaka Math J* 2(1):15–17
- Yarvin N, Rokhlin V (1998) Generalized Gaussian quadratures and singular value decomposition of integral equations. *SIAM J Sci Comput* 20(2):699–718
- Zhou X, Hon YC, Li J (2003) Overlapping domain decomposition method by radial basis functions. *Appl Numer Math* 44:241–255