Advances in Computational Mathematics **21:** 107–125, 2004.

Extremal systems of points and numerical integration on the sphere

Ian H. Sloan and Robert S. Womersley

School of Mathematics, University of New South Wales, Sydney 2052, Australia
E-mail: {I.Sloan,R.Womersley}@unsw.edu.au

Received 29 October 2001; accepted 2 August 2002
Communicated by J. Ward

This paper considers extremal systems of points on the unit sphere $S^r \subseteq \mathbb{R}^{r+1}$, related problems of numerical integration and geometrical properties of extremal systems. Extremal systems are systems of $d_n = \dim \mathbb{P}_n$ points, where $\mathbb{P}_n$ is the space of spherical polynomials of degree at most n, which maximize the determinant of an interpolation matrix. Extremal systems for S^2 of degrees up to 191 (36,864 points) provide well distributed points, and are found to yield interpolatory cubature rules with positive weights. We consider the worst case cubature error in a certain Hilbert space and its relation to a generalized discrepancy. We also consider geometrical properties such as the minimal geodesic distance between points and the mesh norm. The known theoretical properties fall well short of those suggested by the numerical experiments.

Keywords: points on the sphere, extremal systems, cubature rule, worst-case error, generalized discrepancy

AMS subject classification: 65D32, 11K38

1. Introduction

In this paper we consider geometrical properties of point systems on the unit sphere, and related questions of numerical integration, for so-called extremal fundamental systems. These are sets of points $\{x_1, \ldots, x_d\} \subseteq S^r \subseteq \mathbb{R}^{r+1}$ obtained by maximizing the determinant of the interpolation matrix with respect to an arbitrary basis of $\mathbb{P}_n(S^r)$. Here S^r is the unit sphere in $r+1$ dimensions and $\mathbb{P}_n(S^r)$ is the space of spherical polynomials on S^r of degree $\leqslant n$, or equivalently, the restriction to S^r of the polynomials on $\mathbb{R}^{r+1}$ of total degree $\leqslant n$. The requirement on d in order to make the interpolation matrix square is $d = \dim \mathbb{P}_n(S^r)$. For the case $r = 2$, which we shall especially emphasize, we have

$$d = d_n = (n+1)^2$$

(see, for example, [19]). The system of points $\{x_1, \ldots, x_{d_n}\}$ is said to be fundamental if the only polynomial in $\mathbb{P}_n(S^r)$ that vanishes at each point is the zero polynomial, or equivalently, if the determinant of the interpolation matrix is nonzero. Of course it is obvious that a system of points can be fundamental but still of very poor quality.

Our study combines theory and experiment. A general conclusion is that extremal fundamental systems have surprisingly good integration properties if the points are used to determine an interpolatory integration rule and also excellent geometrical properties.

The interpolatory cubature rule associated with the system of points $x_1, \ldots, x_{d_n} \in S^r$ is the rule

$$Q_n(f) := \sum_{j=1}^{d_n} w_j f(x_j) \tag{1.1}$$

obtained by integrating exactly the (spherical) polynomial that interpolates $f \in C(S^r)$ at $x_1, \ldots, x_{d_n}$. For the case $n = 2$ it is known theoretically from the work of Reimer [20] that all of the cubature weights of the rule $Q_n(f)$ are positive, but nothing is known theoretically for larger values of n. In the present study we shall present numerical evidence leading to a conjecture that for $r = 2$ all the cubature weights associated with the interpolatory rule $Q_n(f)$ for an extremal system are positive for *all* n. (We study all values of n up to $n = 50$, then $n = 56, 63, 64, 72, 96, 127, 128$ and 191. For the largest value of n the number of different weights, all positive, is 36,864.)

An interpolatory cubature rule with all weights positive is of intrinsic interest for numerical integration, and as well allows us to obtain an upper bound on the mesh norm.

Positive weight cubature rules for scattered data on spheres have been discussed in [13]; for a survey see [12]. The emphasis in that work is to show that positive weight rules with specified polynomial accuracy can exist even if the choice of the points $x_1, \ldots, x_d$ is not under the control of the user. Our objective is more limited, in that for us the point systems are determined by an algorithm, and hence are fixed. On the other hand, because our rules are interpolatory we have no freedom in choosing the weights, whereas in [13] there is some freedom in the choice of weights, which is exploited there to make the weights positive.

In section 2 we discuss extremal systems, and describe briefly the computations we employ to construct them. Section 3 describes interpolatory cubature, explains the computation of the cubature weights w_j, and discusses the error $I(f) - Q_n(f)$ for interpolatory cubature, where

$$I(f) := \int_{S^r} f(x)\, \mathrm{d}s(x),$$

with $\mathrm{d}s(x)$ denoting surface measure on S^r.

In section 4 we discuss the worst case cubature error for all functions f in the unit ball $B(H)$ of a certain Hilbert space H. The Hilbert space in question has a natural association with the Cui–Freeden generalized discrepancy [3]. Indeed, we shall show that the generalized discrepancy is (apart from a factor of 4π) precisely the worst-case

error, over all $f \in B(H)$, for the special case of an equal-weight (or quasi-Monte Carlo) cubature rule

$$E_n(f) := \frac{4\pi}{d_n} \sum_{j=1}^{d_n} f(x_j). \tag{1.2}$$

Other notions of discrepancy on the sphere have been proposed (see, for example, [8]), of which the most common is the 'spherical cap discrepancy'. However, the generalized discrepancy of [3] has at least one advantage over all other versions we know of, namely that it can be easily computed, even for point systems with many thousands of points.

In our case, where the points $x_1, \ldots, x_{d_n}$ are the points of an extremal system, it seems natural to study the worst-case error also for the interpolatory cubature rule $Q_n(f)$. We establish experimentally that the worst-case errors with $Q_n(f)$ and $E_n(f)$ are very similar for values of n up to $n = 191$.

Some geometrical properties of extremal systems are already known. One such property, which for the case of the circle S^1 traces back to Marcel Riesz [25], is that the minimal geodesic distance between two points of an extremal fundamental system cannot be less than $\pi/(2n)$. The minimal geodesic distance between points is twice the packing radius for identical spherical caps centered at the points. The numerical experiments for $r = 2$ suggest a stronger result, in that for large n the minimal geodesic distance approaches π/n, twice the lower bound.

A second geometrical quantity of interest is the 'mesh norm' of the points, which can be defined as the smallest (spherical) radius of a covering of the sphere by spherical caps centered at $x_1, \ldots, x_d$. We are not aware of any general results for the mesh norm of extremal fundamental systems, but one result follows from a covering theorem of Yudin [31] as extended by Reimer [22], namely that the mesh norm is no larger than $\cos^{-1} z_n$, where z_n is the largest zero of $P_{\lceil n/2 \rceil}$, the Legendre polynomial of degree $\lceil n/2 \rceil$, provided that the interpolatory cubature rule associated with the points $x_1, \ldots, x_d$ has all its weights positive.

The minimum geodesic distance between pairs of points and the mesh norm are discussed in section 5. Finally, in section 6 we give brief conclusions.

2. Extremal systems on S^r

Let $\{\phi_1, \ldots, \phi_{d_n}\}$ be a basis for $\mathbb{P}_n(S^r)$, the space of spherical polynomials on S^r of degree $\leqslant n$. Then the system $\{x_1, \ldots, x_{d_n}\}$ of points on S^r is said to be extremal (see [19]) if it maximizes $|\Delta_n|$, where Δ_n is the determinant

$$\Delta_n(x_1, \ldots, x_{d_n}) = \det\left(\phi_i(x_j)\right)_{i,j=1}^{d_n}.$$

The extremal system is independent of the choice of basis.

The interest in such systems stems from the association with Lagrange interpolation. Let ℓ_j be the fundamental Lagrange polynomial associated with the jth point x_j of the fundamental system $\{x_1, \dots, x_{d_n}\}$, which can be written as

$$\ell_j \in \mathbb{P}_n\big(S^r\big), \quad \ell_j(x_k) = \delta_{jk}, \quad j, k = 1, \dots, d_n.$$

It is easily seen that

$$\ell_j(x) = \frac{\Delta_n(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_{d_n)}}{\Delta_n(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_{d_n})},$$

implying that for an extremal fundamental system we have $|\ell_j(x)| \leqslant 1$ for all $x \in S^r$. The concept appears to go back to Fekete [4] and Auerbach (see [28]), and for the specific case of the sphere to be due to Reimer [19].

Given a function $f \in C(S^r)$, the unique polynomial $\Lambda_n f \in \mathbb{P}_n(S^r)$ that interpolates f at the points of the fundamental system $\{x_1, \dots, x_{d_n}\}$ can be written as

$$\Lambda_n f = \sum_{j=1}^{d_n} f(x_j)\ell_j. \tag{2.1}$$

An immediate consequence is that the norm $\|\Lambda_n\|_\infty$ of the interpolation operator as a map from $C(S^r)$ to $C(S^r)$, given by

$$\|\Lambda_n\|_\infty = \max_{x \in S^r} \sum_{j=1}^{d_n} \big|\ell_j(x)\big|,$$

for the case of an extremal system satisfies $\|\Lambda_n\|_\infty \leqslant d_n$.

2.1. Reproducing kernel basis

In this subsection we restrict attention to $r = 2$. Let $Y_{\ell,k}$ for $\ell = 0, \dots, n$ and $k = 1, \dots, 2\ell + 1$ denote an orthonormal real spherical harmonic basis for $\mathbb{P}_n$ (see [14], for example). The 'reproducing kernel'

$$G_n(x, y) = \sum_{\ell=0}^{n} \sum_{k=1}^{2\ell+1} Y_{\ell,k}(x) Y_{\ell,k}(y), \quad x, y \in S^2, \tag{2.2}$$

has the reproducing kernel property [19]

$$\big(p, G_n(\cdot, x)\big) = p(x) \quad \text{for all } p \in \mathbb{P}_n\big(S^2\big),$$

where (p, q) denotes the inner product

$$(p, q) = \int_{S^2} p(x)q(x)\,\mathrm{d}s(x).$$

Given a fundamental system $\{x_1, \dots, x_{d_n}\}$, the reproducing kernel basis is

$$g_j(x) = G_n(x_j, x), \quad j = 1, \dots, d_n, \ x \in S^2, \tag{2.3}$$

The functions g_j, $j = 1, \ldots, d_n$, belong to $\mathbb{P}_n(S^2)$, and are linearly independent, because the Gram matrix with elements

$$G_{ij} = (g_i, g_j) = G_n(x_i, x_j) \tag{2.4}$$

is nonsingular if $\{x_1, \ldots, x_{d_n}\}$ is a fundamental system.

The addition theorem (see [14])

$$\sum_{k=1}^{2\ell+1} Y_{\ell,k}(x) Y_{\ell,k}(y) = \frac{2\ell+1}{4\pi} P_\ell(x \cdot y), \tag{2.5}$$

where $x \cdot y$ is the usual inner product in $\mathbb{R}^3$, shows that $G_n(x, y)$ is 'bizonal', that is its value depends only on the angle between x and y, and that $G_n(x, y) = \widetilde{G}_n(x \cdot y)$ where

$$\widetilde{G}_n(z) = \frac{1}{4\pi} \sum_{\ell=0}^{n} (2\ell+1) P_\ell(z), \quad z \in [-1, 1], \tag{2.6}$$

where $P_\ell(\cdot)$ is the usual Legendre polynomial. The sum can be written in closed form in terms of Jacobi polynomials (see [7]), but can also be conveniently computed from (2.6) by upward recurrence of the Legendre polynomials.

The value of $\widetilde{G}_n(1)$ is [19],

$$G_n(x, x) = \widetilde{G}_n(1) = \frac{d_n}{4\pi} \quad \forall x \in S^2. \tag{2.7}$$

Thus the matrix G in (2.4) has equal diagonal elements, $G_{ii} = d_n/(4\pi)$ for $i = 1, \ldots, d_n$, and

$$\text{trace}(G) = \sum_{j=1}^{d_n} \lambda_j = \sum_{j=1}^{d_n} G_{jj} = \frac{d_n^2}{4\pi},$$

where $\lambda_1, \ldots, \lambda_{d_n}$ are the eigenvalues of G. The mean eigenvalue of G is $\lambda_{\text{avg}} = d_n/(4\pi)$. An upper bound on the determinant of G is achieved by $G = \lambda_{\text{avg}} I$, giving

$$\det(G) \leqslant \left(\frac{d_n}{4\pi}\right)^{d_n}. \tag{2.8}$$

A lower bound is available from the work of Reimer [21], who computed the average of $\det(G)$ over all choices of the points $x_1, \ldots, x_{d_n}$. Since our extremal systems must give a better than average value of the determinant, this gives a lower bound of

$$\frac{d_n!}{(4\pi)^{d_n}} \leqslant \det(G), \tag{2.9}$$

where the quantity on the left is Reimer's average. It has been observed, see [30], that calculated values of $\det(G)$ for extremal systems are much closer to the upper than the lower bounds.

2.2. Calculating extremal systems

Reimer and Sündermann [24] proposed a Remez-like algorithm based on finding $\max_{x\in S^2} \ell_j(x)$ for calculating extremal systems, and give results for $n = 1, \ldots, 6$. For the present computation of extremal systems it seems convenient, following [30], to proceed in a different way. It follows from (2.2) that the matrix G defined by (2.4) can be written as $G = A^{\mathrm{T}}A$, where A is the basis matrix obtained from the spherical harmonic basis, see [30]. Clearly G is positive semi-definite for any point set, and $\det(G) = \det(A)^2 \geqslant 0$, so an extremal system can be obtained by maximizing the determinant of the symmetric positive definite matrix G.

Instead of maximizing $\det(G)$ we maximize

$$\psi(x_1, \ldots, x_{d_n}) = \log\det(G), \tag{2.10}$$

where log denotes the natural logarithm, in this way avoiding problems of overflow. As G is rotationally invariant, the first point x_1 is conveniently fixed at the north pole and the second point x_2 on the prime meridian, as in [5,30]. A spherical parametrization $\theta_j \in [0, \pi]$ and $\phi_j \in [0, 2\pi)$ of the points x_j, $j = 1, \ldots, d_n$, then has $2d_n - 3$ variables. The gradient of ψ can be calculated through the use of

$$\nabla_{x_j}\psi(x_1, \ldots, x_{d_n}) = 2\sum_{k=1}^{d_n} x_k G^{-1}_{kj}\dot{G}_{kj},$$

where $\dot{G}_{kj} = \widetilde{G}'_n(x_k \cdot x_j)$. The function $\widetilde{G}_n(z)$ in (2.6) is only defined for $z \in [-1, 1]$, so the points x_j must lie on the unit sphere. Optimization methods for problems with nonlinear constraints (such as $x_j \cdot x_j = 1$) often make better progress by allowing variables to move off the constraints, but that is not allowable in this case. Working with the spherical parametrization of the points x_j, $j = 1, \ldots, d_n$, avoids this difficulty. An unconstrained problem is produced by ignoring the simple bounds $\theta_j \in [0, \pi]$ and $\phi_j \in [0, 2\pi)$ and using periodicity to map the spherical parameters back to these ranges. Care must be exercised in doing this when using an optimization method that updates information based on differences in iterates (for example, quasi-Newton methods).

Both a restarted conjugate gradient algorithm, followed by a Newton method using a finite difference approximation to the Hessian and a limited memory BFGS method ([15] for example) were used to maximize ψ. This only finds a local maximizer. The algorithm is then tried from several different starting points including the equal area points [18], the generalized spiral points [26], the minimum energy points of Fliege and Maier [5], the spherical designs of Hardin and Sloane [9,10], and randomly generated points, and the best final value selected. We have observed many different local maxima with relatively little difference in the values of ψ. Deterministic methods for global optimization, for example, [11], can be tried for small numbers of points, as in [27], but are not realistic for large numbers of points.

The calculated values of $\log\det(G)$ for a typical sample of degrees, along with the lower and upper bounds from (2.8) and (2.9), are given in table 1. The column labelled

Table 1
Log det(G) for some of the computed extremal systems, and bounds.

n	d_n	$\log(d_n!/(4\pi)^{d_n})$ lower bound	R & S	$\log\det(G)$	$\log((d_n/4\pi)^{d_n})$ upper bound
1	4	−6.95	−4.58	−4.58	−4.58
2	9	−9.98		−3.21	−3.00
3	16	−9.82	3.39	3.39	3.87
4	25	−5.27	16.00	16.14	17.20
5	36	4.60	35.96	36.17	37.89
6	49	20.55	63.39	64.09	66.68
7	64	43.18		100.69	104.18
8	81	73.05		146.19	150.94
9	100	110.64		201.56	207.41
10	121	156.35		266.32	274.04
16	289	620.88		885.26	906.13
32	1089	3774.52		4768.27	4859.11
48	2401	10215.34		12399.56	12611.53
64	4225	20360.09		24203.22	24579.99
96	9409	52869.00		61398.34	62272.50
128	16641	102990.28		118065.05	119625.50
191	36864	257463.11		290802.67	294320.94

R & S are the values for the point systems reported by Reimer and Sündermann [24]. The present values are all at least as large, and therefore at least as good as those calculated by Reimer and Sündermann, and are better for $n = 4, 5, 6$.

Figure 1 gives a picture of the calculated point set for $n = 64$, with $d_n = 4225$ points. A Delaunay triangulation (see [2,16]) has been added to make the picture easier to comprehend. Clearly the triangles are remarkably regular in size and shape.

The computed points and cubature weights of the (approximate) extremal systems are given at the web site `http://www.maths.unsw.edu.au/~rsw/Sphere` for all the values of n from 1 to 51, then for $n = 56, 63, 64, 72, 96, 127, 128$, and 191. It should be noted that the computations for the larger values of n are extremely demanding on computer resources, especially memory. For the largest value of n, $n = 191$, the matrix G is a dense 36,864 by 36,864 symmetric matrix, whose determinant is optimized as a function of more than 73,000 variables. Only very limited attempts have been made to find the global maximum for the larger values of n. The calculated values always provide lower bounds on the true global maxima.

3. Interpolatory cubature

The interpolatory cubature rule corresponding to the fundamental system $\{x_1, \ldots, x_{d_n}\} \subseteq S^r$ can be written

$$Q_n(f) = \int_{S^r} (\Lambda_n f)(x)\, \mathrm{d}s(x) = \sum_{j=1}^{d_n} w_j f(x_j),$$

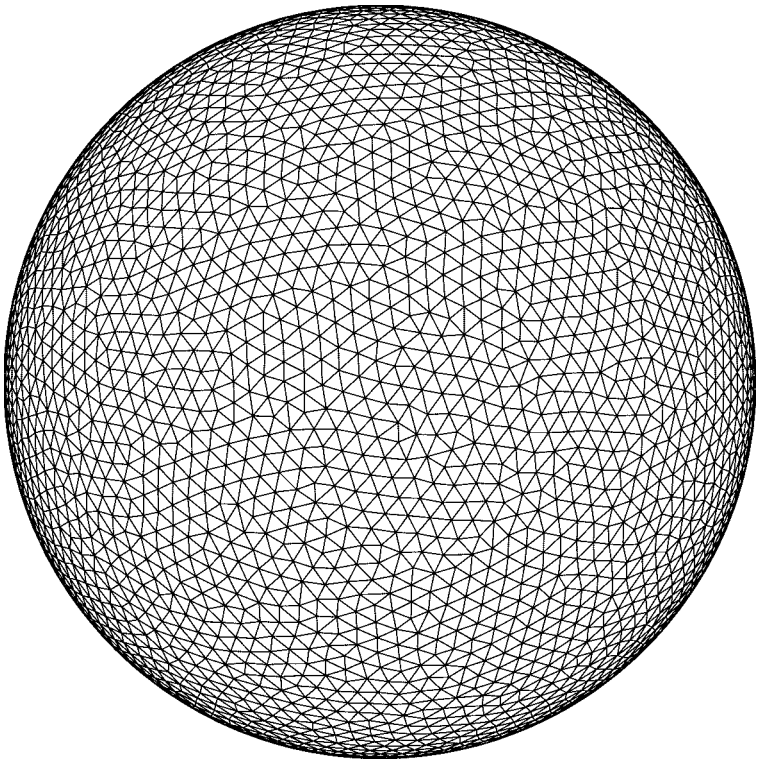

Figure 1. Computed extremal system for $n = 64$, $d_n = 4225$.

where from (2.1)

$$w_j = \int_{S^r} \ell_j(x)\, \mathrm{d}s(x).$$

The practical computation of the weights proceeds differently. The reproducing kernel basis functions $g_j(x) = G_n(x_j, x)$, where $G_n(x, y)$ is the reproducing kernel for $\mathbb{P}_n(S^r)$ [19], satisfy

$$\int_{S^r} g_j(x)\, \mathrm{d}s(x) = 1 \quad \text{for } j = 1, \ldots, d_n,$$

so the condition that the cubature rule (1.1) is exact for all polynomials in $\mathbb{P}_n$ can be written as

$$Gw = e, \tag{3.1}$$

where w is the vector of cubature weights, e is the vector of 1's in $\mathbb{R}^{d_n}$ and G is the matrix defined in (2.4). In particular the cubature rule is exact for the constant polynomial $1 \in \mathbb{P}_n$, so $\sum_{j=1}^{d_n} w_j = |S^r|$. Hence the average weight is $w_{\text{avg}} = |S^r|/d_n = 1/\lambda_{\text{avg}}$, or $w_{\text{avg}} = 4\pi/d_n$ for $r = 2$. The computed minimum and maximum values of the ratio of the cubature weights to the average weight for the extremal systems of section 2 are

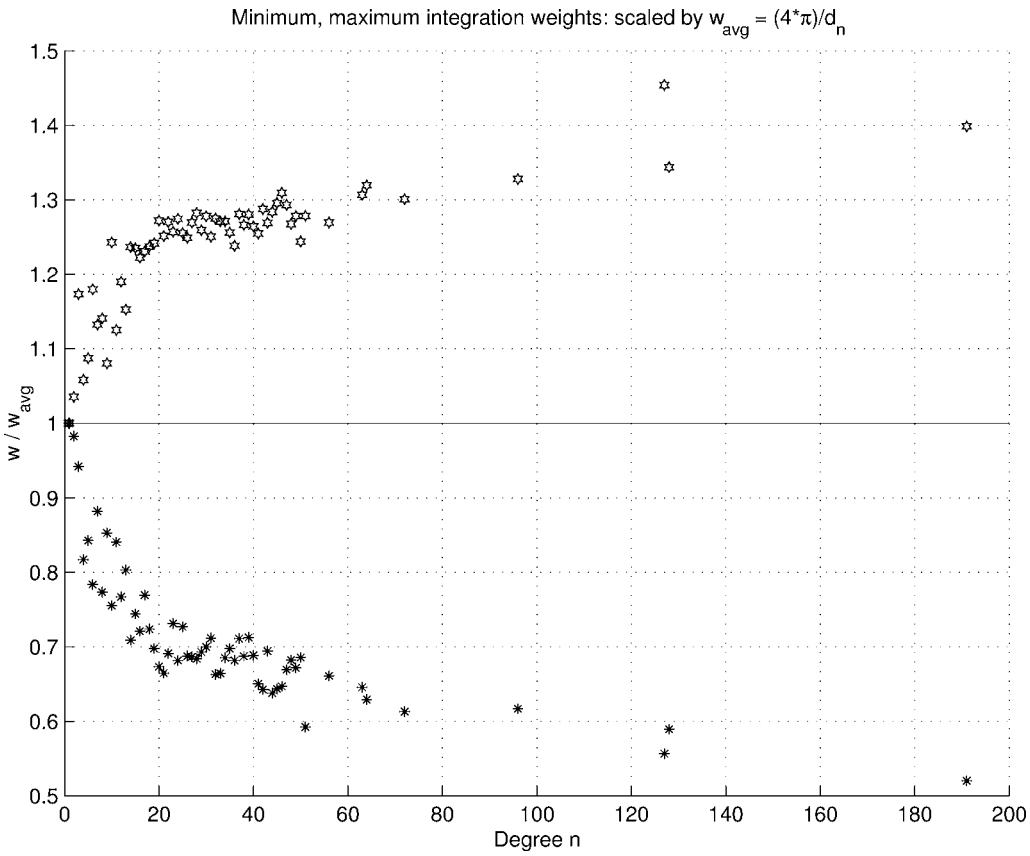

Figure 2. Scaled minimum and maximum cubature weights for the computed extremal systems.

plotted in figure 2. It is notable that all the computed weights lie between $1/2$ and $3/2$ of the mean weight.

It should be said that for other point systems and for the larger values of n in table 1 it is often impossible to solve for the weights using (3.1), because of ill-conditioning of the matrix G. It is a feature of extremal systems that the matrix G is remarkably well conditioned. For example, the 1-norm condition numbers $\kappa_1(G)$ of the final matrix G for $n = 64, 96, 128$ and 191 respectively are only 3.2×10^3, 1.4×10^4, 2.8×10^4 and 1.2×10^5.

We have already noted that the average eigenvalue of G is given by $\lambda_{\text{avg}} = d_n/4\pi$, and so is fixed in advance for any point system. The process of maximizing the log of the determinant of G (and hence of maximizing the determinant itself) can therefore be interpreted as trying to make all the eigenvalues as close as possible to the average eigenvalue. In the case $n = 1$ this can be achieved: for in this case the extremal system is a regular tetrahedron, and the matrix G is diagonal. It is known that for $r \geqslant 2$ and $n \geqslant 3$ that the matrix G cannot be diagonal [19], since otherwise one would have a contradiction to a famous theorem of [1] on the non-existence of tight spherical designs.

The error for an interpolatory cubature rule is

$$I(f) - Q_n(f) = \int_{S^r} (f - \Lambda_n f)\,\mathrm{d}s(x),$$

where $\Lambda_n f$ is the interpolant of f. Let

$$\mathcal{E}_n(f) = \inf_{p\in\mathbb{P}_n(S^r)} \|f - p\|_\infty$$

denote the error of best uniform approximation. For any polynomial p in $\mathbb{P}_n(S^r)$ we see that

$$\|f - \Lambda_n f\|_\infty \leqslant \big\|\big(\mathcal{I} - \Lambda_n\big)(f - p)\big\|_\infty \leqslant \big(1 + \|\Lambda_n\|_\infty\big)\mathcal{E}_n(f),$$

where $\mathcal{I}$ is the identity operator. This gives a bound

$$\big|I(f) - Q_n(f)\big| \leqslant \big|S^r\big|\big(1 + \|\Lambda_n\|_\infty\big)\mathcal{E}_n(f). \tag{3.2}$$

The growth of the interpolation operator norm $\|\Lambda_n\|_\infty$ was studied in [30]. The numerical experiments suggest that the growth in $\|\Lambda_n\|_\infty$ is much less than the upper bound $d_n = (n+1)^2$ for n up to 50. Estimates of $\mathcal{E}_n(f)$ are available from the so-called Jackson theorems for the sphere, see [17].

A better bound follows from the property that $Q_n(f)$ is exact for all polynomials of degree $\leqslant n$:

$$\big|I(f) - Q_n(f)\big| = \big|(I - Q_n)(f - p)\big| \leqslant \big(\|I\| + \|Q_n\|\big)\|f - p\|_\infty,$$

giving

$$\big|I(f) - Q_n(f)\big| \leqslant \left(\big|S^r\big| + \sum_{j=1}^{d_n} |w_j|\right)\mathcal{E}_n(f),$$

or if the weights are all positive

$$\big|I(f) - Q_n(f)\big| \leqslant 2\big|S^r\big|\mathcal{E}_n(f).$$

4. Worst-case cubature error bound in a Hilbert space

The most convenient version of the generalized discrepancy of a set of points $x_1, \ldots, x_d \in S^2$ has the explicit form

$$D(x_1, \ldots, x_d) = \frac{1}{2\sqrt{\pi}\,d}\left[\sum_{j=1}^{d}\sum_{k=1}^{d}\left(1 - 2\log\left(1 + \sqrt{\frac{1 - x_j\cdot x_k}{2}}\right)\right)\right]^{1/2} \tag{4.1}$$

given by Cui and Freeden [3]. We show that $D(x_1, \ldots, x_d)$ has a natural interpretation, as the worst-case error (apart from a constant factor) for the equal-weight cubature rule (1.2) applied to a function $f \in B(H)$, where $B(H)$ is the unit ball in a certain Hilbert space H.

In this section we define the Hilbert space H, and show how the worst-case error can be computed for any cubature rule (1.1) with $x_1, \ldots, x_d \in S^r$. At the end we establish the connection with the generalized discrepancy (4.1).

We fix $r = 2$, and let $B_0, B_1, \ldots$ be a sequence of positive real numbers with the property that there exist constants $c_1, c_2 > 0$ such that

$$B_0 = 1, \qquad c_1 \ell^3 \leqslant B_\ell \leqslant c_2 \ell^3 \quad \text{for } \ell \geqslant 1. \tag{4.2}$$

Then for $f, g \in C^\infty(S^2)$ we may define the inner product

$$(f, g)_H = \sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} B_\ell \hat{f}_{\ell k} \hat{g}_{\ell k}, \tag{4.3}$$

and norm

$$\|f\|_H = (f, f)_H^{1/2} = \left(\sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} B_\ell \hat{f}_{\ell k}^2 \right)^{1/2},$$

where

$$\hat{f}_{\ell k} = \int_{S^2} f(x) Y_{\ell,k}(x) \, \mathrm{d}s(x).$$

The Hilbert space H may then be defined as the completion of $C^\infty(S^2)$ with respect to this inner product. The norm $\|f\|_H$ is clearly equivalent to the $H^{3/2}(S^2)$ norm, where the $H^s(S^2)$ norm is defined (following [6]) by

$$\|f\|_s = \left(\sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} \left(\ell + \frac{1}{2} \right)^{2s} \hat{f}_{\ell k}^2 \right)^{1/2}.$$

Later we shall make a specific choice for B_ℓ, following [3], to simplify the calculations.

Associated with the inner product (4.3) and the Hilbert space H is the kernel

$$K(x, y) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} \frac{1}{B_\ell} Y_{\ell,k}(x) Y_{\ell,k}(y) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi B_\ell} P_\ell(x \cdot y), \tag{4.4}$$

where in the last step we used the addition theorem (2.5). Note that the assumption (4.2) ensures that the sum in (4.4) exists for all $x, y \in S^2$.

The kernel $K(x, y)$ is in fact the reproducing kernel in the Hilbert space H, that is to say, for each $f \in H$ and each $y \in S^2$ we have

$$K(\cdot, y) \in H, \quad \big(f, K(\cdot, y)\big)_H = f(y). \tag{4.5}$$

The latter property follows because by definition the left side of this equation is

$$\sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} B_\ell \hat{f}_{\ell k} \frac{Y_{\ell,k}(y)}{B_\ell} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} \hat{f}_{\ell k} Y_{\ell,k}(y) = f(y).$$

For a given $f \in H$, the reproducing kernel property (4.5) allows us to express the error for a given cubature rule $Q_n(f)$ as

$$I(f) - Q_n(f) = (f, \xi)_H,$$

where ξ (the 'representer' of the error) is given by

$$\xi(x) = \int_{S^2} K(x, y)\, \mathrm{d}s(y) - \sum_{j=1}^{d_n} w_j K(x, x_j). \tag{4.6}$$

In turn it follows that the worst-case error for all f in $B(H)$ is

$$\begin{aligned} e(Q_n) &:= \sup\left\{|I(f) - Q_n(f)|\colon \|f\|_H \leqslant 1\right\} \\ &= \sup\left\{|(f, \xi)_H| : \|f\|_H \leqslant 1\right\} \\ &= \|\xi\|_H \\ &= (\xi, \xi)_H^{1/2} \\ &= \left[\int_{S^2}\left(\int_{S^2} K(x, x')\, \mathrm{d}s(x)\right)\mathrm{d}s(x') - 2\sum_{j=1}^{d_n} w_j \int_{S^2} K(x, x_j)\, \mathrm{d}s(x)\right. \\ &\qquad \left. + \sum_{j'=1}^{d_n}\sum_{j=1}^{d_n} w_j w_{j'} K(x_j, x_{j'})\right]^{1/2}, \end{aligned}$$

where in the last step we used (4.6) and the reproducing kernel property (4.5). The last expression can be simplified, using

$$\int_{S^2} K(x, x')\, \mathrm{d}s(x) = 1, \quad \text{for } x' \in S^2,$$

giving, with $\sum_{j=1}^{d_n} w_j = 4\pi$,

$$e(Q_n) = \left[-4\pi + \sum_{j'=1}^{d_n}\sum_{j=1}^{d_n} w_j w_{j'} K(x_j, x_{j'})\right]^{1/2}. \tag{4.7}$$

Following [3], we now make a special choice for the sequence (B_ℓ), namely

$$B_\ell = \begin{cases} 1 & \text{if } \ell = 0, \\ (2\ell + 1)\ell(\ell + 1) & \text{for } \ell = 1, 2, \ldots, \end{cases} \tag{4.8}$$

for the very good reason that this choice leads to a closed form expression for the reproducing kernel K: for (4.4) now gives

$$K(x, y) = \frac{1}{4\pi} + \frac{1}{4\pi}\sum_{\ell=1}^{\infty} \frac{P_\ell(x \cdot y)}{\ell(\ell + 1)},$$

which with the identity

$$\sum_{\ell=1}^{\infty} \frac{P_\ell(z)}{\ell(\ell+1)} = 1 - 2\log\left(1 + \sqrt{\frac{1-z}{2}}\right)$$

becomes

$$K(x, y) = \frac{1}{2\pi}\left(1 - \log\left(1 + \sqrt{\frac{1 - x \cdot y}{2}}\right)\right).$$

With this simple formula, the worst-case error $e(Q_n)$ for the cubature rule Q_n can be computed easily from (4.7).

For the special case of the equal weight rule given by (1.2) it is easily checked that the worst-case error is related to the generalized discrepancy by

$$e(E_n) = 4\pi D(x_1, \ldots, x_d). \tag{4.9}$$

A theoretical bound on $e(Q_n)$ is easily obtained for any rule Q_n which is exact for polynomials of degree $\leqslant n$, and which has positive weights.

Theorem 4.1. For $n \geqslant 1$, let Q_n be a positive weight cubature rule on S^2 which integrates exactly all $p \in \mathbb{P}_n$. With B_ℓ defined by (4.8),

$$e(Q_n) \leqslant \frac{\sqrt{4\pi}}{(n+1)^{1/2}} = \frac{\sqrt{4\pi}}{d_n^{1/4}}.$$

Proof. From (4.7) and (4.4) we have

$$e(Q_n)^2 = -4\pi + \sum_{\ell=0}^{\infty} \sum_{k=1}^{2\ell+1} \frac{1}{B_\ell} \left(\sum_{j=1}^{d_n} w_j Y_{\ell,k}(x_j) \right)^2.$$

Now for $\ell \leqslant n$ we have

$$\sum_{j=1}^{d_n} w_j Y_{\ell,k}(x_j) = Q_n(Y_{\ell,k}) = I(Y_{\ell,k}) = \delta_{\ell 0} I(Y_{0,k}) = \sqrt{4\pi}\,\delta_{\ell 0},$$

while for $\ell \geqslant n+1$ the addition theorem (2.5) and $P_\ell(x \cdot y) \leqslant 1\ \forall x, y$ give

$$\begin{aligned} \sum_{k=1}^{2\ell+1} \left(\sum_{j=1}^{d_n} w_j Y_{\ell,k}(x_j) \right)^2 &= \sum_{j=1}^{d_n} w_j \sum_{j'=1}^{d_n} w_{j'} \frac{2\ell+1}{4\pi} P_\ell(x_j \cdot x_{j'}) \\ &\leqslant (2\ell+1)4\pi. \end{aligned}$$

We therefore obtain, using (4.8),

$$e(Q_n)^2 \leqslant \sum_{\ell=n+1}^{\infty} \frac{(2\ell+1)4\pi}{B_\ell} = 4\pi \sum_{\ell=n+1}^{\infty} \frac{1}{\ell(\ell+1)} = \frac{4\pi}{n+1}.$$

□

Table 2
Worst case errors for the interpolatory rule Q_n and equal-weight rule E_n, and generalized discrepancy for some computed extremal systems.

n	d_n	$e(Q_n)$	$e(E_n)$	Cui and Freeden discrepancy D
1	4	1.146686	1.146686	0.09125
2	9	0.620391	0.619657	0.04931
4	25	0.287603	0.287061	0.02284
8	81	0.118700	0.118600	0.00944
16	289	0.045816	0.045690	0.00364
32	1089	0.016924	0.016890	0.00134
48	2401	0.009360	0.009347	0.00074
64	4225	0.006129	0.006118	0.00049
96	9409	0.003361	0.003361	0.00027
128	16641	0.002194	0.002190	0.00017
191	36864	0.001210	0.001208	0.00009

Table 2 gives, for typical n, the worst-case errors for the interpolatory rule Q_n and the equal-weight rule E_n. Clearly there is very little difference. It also gives the Cui and Freeden generalized discrepancy D for the computed extremal systems, related to $e(E_n)$ by (4.9). These values are better than all the values reported by Cui and Freeden for values of d up to about 450 and for a variety of point systems. They are close to, but still better than, the best results in [3], which are those for a Hammersley point sequence with respect to $\cos\theta$ and ϕ.

The present discussion allows us to express these conclusions in terms of worst-case errors: the interpolatory cubature rules based on extremal systems have smaller worst-case errors in $B(H)$ than all of the quasi-Monte Carlo rules considered in [3].

The error bound in theorem 4.1 can be used for the interpolatory cubature rule based on the extremal systems of section 2, because all the weights w_j were found to be positive. On the other hand, the bound in theorem 4.1 greatly overestimates the error: it gives a bound of order $O(d_n^{-1/4})$, whereas a least squares fit accurately approximates $e(Q_n)$ by $3.217 d_n^{-0.7503}$, with an exponent of approximately $-3/4$.

5. Geometric properties

This section discusses two geometric properties of point systems on the sphere, namely the minimum geodesic distance between points (twice the packing radius) and the mesh norm (covering radius).

5.1. Minimum geodesic distance between points

The geodesic distance between points x and y on S^r is $\mathrm{dist}(x, y) = \cos^{-1}(x \cdot y)$, where $x \cdot y$ is the usual inner product in $\mathbb{R}^{r+1}$.

Reimer [19] has established the following result for the minimum distance between pairs of points of an extremal system.

Theorem 5.1. For $n \geqslant 1$, let $\{x_1, \ldots, x_{d_n}\} \subseteq S^r$ be the points of an extremal system for $\mathbb{P}_n(S^r)$. Then

$$\cos^{-1}(x_j \cdot x_k) \geqslant \frac{\pi}{2n} \quad \text{for } j \neq k.$$

The proof proceeds by observing that ℓ_j (the fundamental Lagrange polynomial associated with x_j), when restricted to the great circle defined by x_j and x_k (on the assumption that $x_j \cdot x_k \neq -1$), is a trigonometric polynomial of degree n, which takes a maximum at x_j (since $\ell_j(x_j) = 1$ and $|\ell_j(x)| \leqslant 1$ for all $x \in S^r$) and vanishes at x_k. The result then follows from an analogous result by Marcel Riesz [25] for trigonometric polynomials of degree $\leqslant n$ on the circle S^1.

In figure 3 we show the minimum geodesic distance between pairs of points in our extremal systems as a function of n. For $n = 1$ the points of the regular tetrahedron have

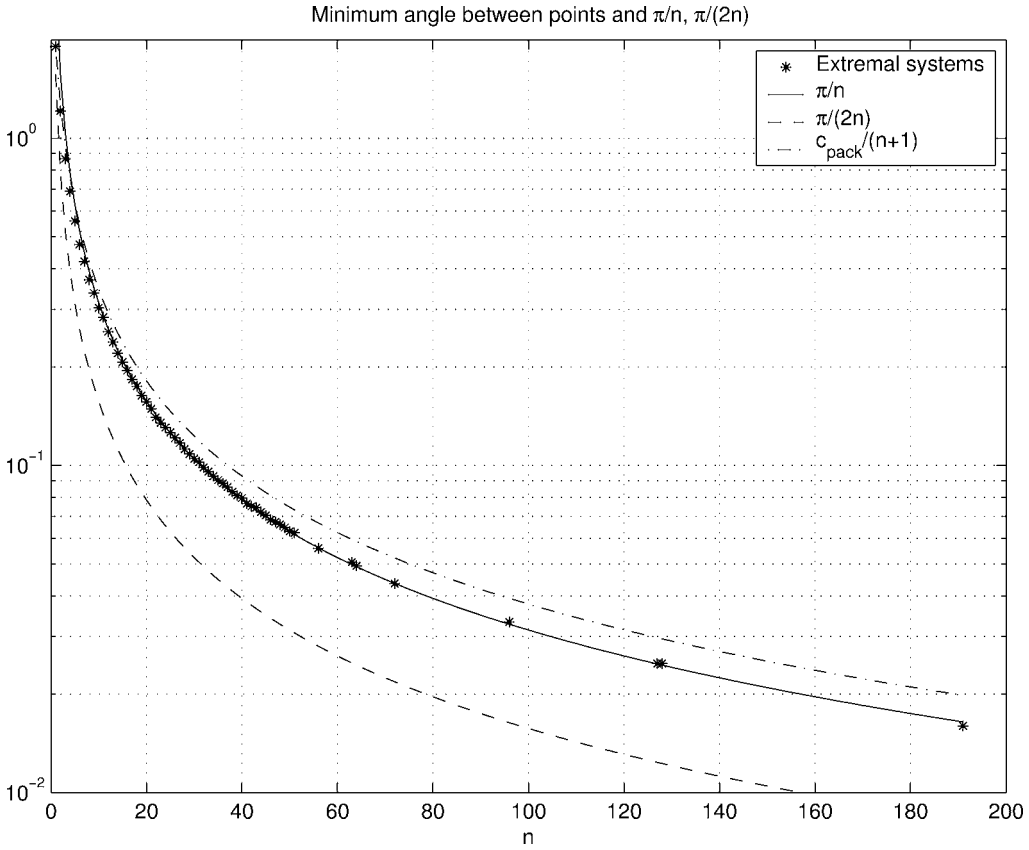

Figure 3. Minimum angle between points for the computed extremal systems.

$x_j \cdot x_k = -1/3$ for $j \neq k$ or $\cos^{-1}(x_j \cdot x_k) \approx \pi/1.644$. We see that the theorem holds without much room to spare for small values of n, but that as n increases the minimum geodesic distance between pairs appears to approach π/n. We have no explanation of this result, but perhaps the following observation gives some insight. The proof in effect makes use of the fact that the most extreme behaviour of a trigonometric polynomial of degree n or less on the great circle defined by x_j and x_k is that exhibited by the restriction of $\cos(n(\cos^{-1}(x \cdot x_j)))$ to the great circle. However, this trigonometric polynomial does not at all have the character of the fundamental Lagrange (trigonometric) polynomial for the circle S^1 with equally spaced interpolation points, which is given by

$$L_j(\theta) = \frac{\sin(n + 1/2)\theta}{(2n + 1)\sin(1/2)\theta},$$

and which has its first zero at $\theta = \cos^{-1}(x \cdot x_j) = \pi/n$. The numerical results in figure 3 seem to suggest that for large n the fundamental Lagrange polynomial ℓ_j, when restricted to the great circle defined by x_j and x_k, behaves for x near x_j more like $L_j(\theta)$ than like $\cos(n(\theta))$.

5.2. *Mesh norm*

The 'mesh norm' of a set of points $x_1, \ldots, x_d \in S^r$ may be defined as

$$h(x_1, \ldots, x_d) := \max_{x \in S^r} \min_{j=1,\ldots,d} \cos^{-1}(x \cdot x_j).$$

In words, it is the largest geodesic distance from a point $x \in S^r$ to the nearest mesh point x_j. An equivalent definition is that $h(x_1, \ldots, x_d)$ is the minimum (spherical) radius of a covering of S^r by spherical caps of equal spherical radius centred at $x_1, \ldots, x_d$.

The mesh norms for the calculated extremal systems are plotted in figure 4.

We are not aware of any theoretical results for the mesh norms of extremal fundamental systems, but recently Reimer [22, theorem 4.1] obtained an upper bound on the mesh norm of any system associated with a positive-weight cubature rule which integrates exactly all polynomials up to some specified degree. Reimer's result is applicable to all of the extremal fundamental systems reported in section 2, because in every case the interpolatory cubature rule associated with these systems (which integrates exactly all polynomials of degree up to n) was found to have positive weights. Reimer's result, adapted to the present circumstances, is as follows.

Theorem 5.2. For $n \geqslant 1$, let $x_1, \ldots, x_{d_n}$ be a fundamental system of points on S^2 such that the associated interpolatory cubature rule has all its weights positive. Then

$$h(x_1, \ldots, x_{d_n}) \leqslant \cos^{-1} z_n,$$

where z_n is the largest zero of $P_{\lceil n/2 \rceil}$.

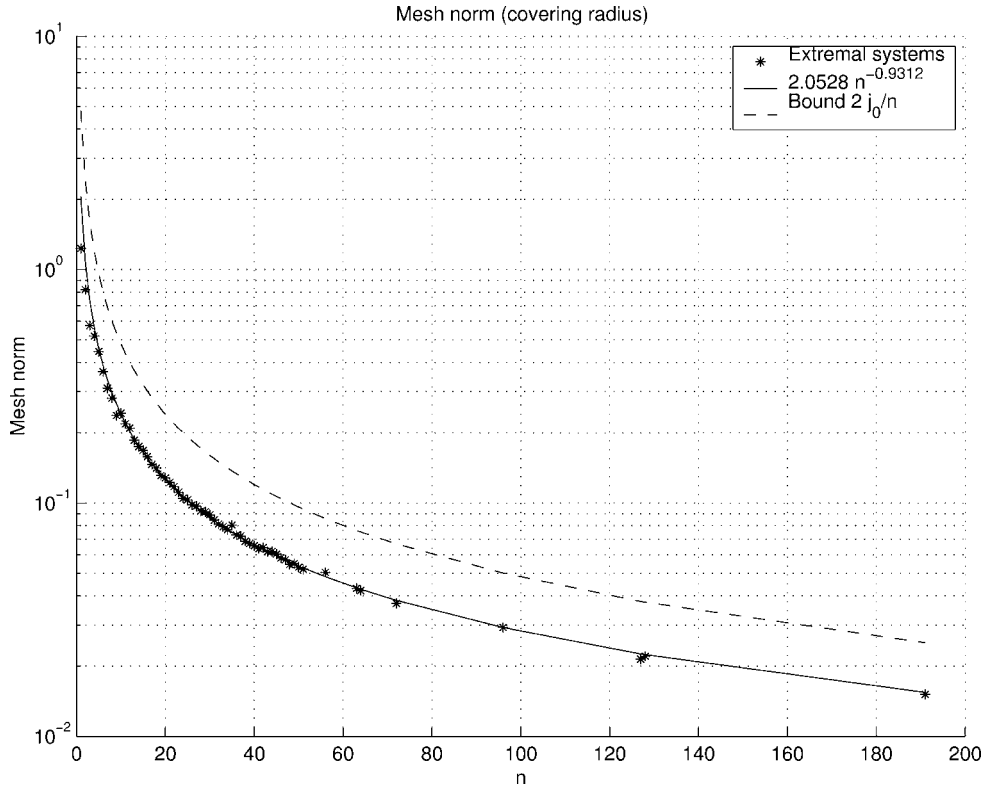

Figure 4. Mesh norm $h(x_1, \ldots, x_d)$ for the computed extremal systems.

Reimer's result is a slight generalization of a result of Yudin [31] for spherical designs, which may be defined as point systems for which equal-weight cubature rules integrate exactly all spherical polynomials up to some degree.

Reimer [23] notes the result [29, p. 186] that the largest zero of P_μ behaves asymptotically like $\cos j_0/\mu$, where j_0 is the smallest positive zero of the Bessel function J_0. This gives $h(x_1, \ldots, x_d) \sim 2j_0/n \approx 4.8097/n$. From figure 4 we observe that this bound overestimates the true mesh norm by something less than a factor of 2.

6. Conclusions

The extremal systems computed here, for n as large as 191, turn out to have excellent geometrical properties. The associated interpolatory integration rules have all weights positive and good approximation properties. Some properties are known theoretically (such as the positivity of the weights for the case $n = 2$), but many properties suggested by the numerical experiments are not known. The most current points and weights are available from `http://www.maths.unsw.edu.au/~rsw/Sphere`. Further research on extremal systems is clearly needed.

Acknowledgements

The support of the Australian Research Council is acknowledged. We are also grateful to W. Freeden for stimulating our interest in interpolatory cubature for large values of n. The larger computations were performed on the NEC SX-5 and SGI Origin 2400 of the Australian Centre for Advanced Computing and Communication (ac3), and the Compaq SC of the Australian Partnership for Advanced Computing (APAC). The programming assistance of Paul Leopardi is also gratefully acknowledged.

References

[1] E. Bannai and R.M. Damerell, Tight spherical design I, Math. Soc. Japan 31 (1979) 199–207.
[2] J.H. Conway and N.J.A. Sloane, *Sphere Packings: Lattices and Groups*, 3rd ed. (Springer, New York/Berlin, 1999).
[3] J. Cui and W. Freeden, Equidistribution on the sphere, SIAM J. Sci. Comput. 18 (1997) 595–609.
[4] M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, Math. Z. 17 (1923) 228–249.
[5] J. Fliege and U. Maier, The distribution of points on the sphere and corresponding cubature formulae, IMA J. Numer. Anal. 19 (1999) 317–334; `http://www.mathematik.uni-dortmund.de/lsx/fliege/nodes.html`.
[6] W. Freeden, T. Gervens and M. Schreiner, *Constructive Approximation on the Sphere* (Clarendon Press, Oxford, 1998).
[7] M. Ganesh, I. Graham and J. Sivaloganathan, A pseudospectral three-dimensional boundary integral method applied to a nonlinear model problem from finite elasticity, SIAM J. Numer. Anal. 31 (1994) 1378–1414.
[8] P.J. Grabner, B. Klinger and R. F. Tichy, Discrepancies of point sequences on the sphere and numerical integration, in: *Multivariate Approximation: Recent Trends and Results*, eds. W. Haußmann, K. Jetter and M. Reimer, Mathematical Research, Vol. 101 (Academie-Verlag, Berlin, 1997) pp. 95–112.
[9] R.H. Hardin and N.J.A. Sloane, Spherical designs, ATT Research Laboratory, `http://www.research.att.com/~njas/sphdesigns`.
[10] R.H. Hardin and N.J.A. Sloane, McLaren's improved snub cube and other new spherical designs in three dimensions, Discrete Comput. Geom. 15 (1996) 429–441.
[11] R. Horst and H. Tuy, *Global Optimizations – Deterministic Approaches* (Springer, Berlin, 1996).
[12] H.N. Mhaskar, F.J. Narcowich and J.D. Ward, Representing and analyzing scattered data on spheres, in: *Multivariate Approximation and Applications*, eds. D.L.D. Leviaton and A. Pinkus (Cambridge Univ. Press, Cambrige, 2001) pp. 44–72.
[13] H.N. Mhaskar, F.J. Narcowich and J.D. Ward, Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature, Math. Comp. 70 (2001) 1113–1130.
[14] C. Müller, *Spherical Harmonics*, Lecture Notes in Mathematics, Vol. 17 (Springer, Berlin/New York, 1966).
[15] J. Nocedal and S.J. Wright, *Numerical Optimization* (Springer, New York, 1999).
[16] A. Okabe, B. Boots and K. Sugihara, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams* (Wiley, Chichester/New York, 1992).
[17] D.L. Ragozin, Constructive polynomial approximation on sphere and projective spaces, Trans. Amer. Math. Soc. 162 (1971) 157–170.
[18] E.A. Rahkmanov, E.B. Saff and Y.M. Zhou, Minimal discrete energy on the sphere, Math. Res. Lett. 1 (1994) 647–662.
[19] M. Reimer, *Constructive Theory of Multivariate Functions* (BI Wissenschaftsverlag, Mannheim, 1990).

[20] M. Reimer, Quadrature rules for the surface integral of the unit sphere based on extremal fundamental systems, Math. Nachrichten 169 (1994) 235–241.
[21] M. Reimer, The average size of certain Gram-determinants and interpolation on non-compact sets, in: *Multivariate Approximation and Splines*, Mannheim, 1996, International Series of Numerical Mathematics, Vol. 125 (Birkhäuser, Basel, 1998) pp. 235–244.
[22] M. Reimer, Spherical polynomial approximation: A survey, in: *Advances in Multivariate Approximation*, eds. W. Haußmann, K. Jetter and M. Reimer (Wiley, Berlin, 1999) pp. 231–252.
[23] M. Reimer, Hyperinterpolation on the sphere at the minimal projection order, J. Approx. Theory 104 (2000) 272–286.
[24] M. Reimer and B. Sündermann, A Remez-type algorithm for the calculation of extremal fundamental systems for polynomial spaces on the sphere, Computing 37 (1986) 43–58.
[25] M. Riesz, Eine trigonometrische Interpolationsformel und einige Ungleichungen für polynome, Jahresbericht Deutsches Mathematische Vereinigung 23 (1914) 354–368.
[26] E.B. Saff and A.B.J. Kuijlaars, Distributing many points on a sphere, Math. Intelligencer 19 (1997) 5–11.
[27] W.J.H. Stortelder, J.J.B. Swart and J.D. Pinter, Finding elliptic fekete point sets: Two numerical solution approaches, J. Comput. Appl. Math. 130 (2001) 205–216.
[28] A.E. Taylor, A geometric theorem and its application to biorthogonal systems, Bull. Amer. Math. Soc. 53 (1947) 614–616.
[29] F.G. Tricomi, *Vorlesungen über Orthogonalreihen* (Springer, Berlin, 1955).
[30] R.S. Womersley and I.H. Sloan, How good can polynomial interpolation on the sphere be?, Adv. Comput. Math. 14 (2001) 195–226.
[31] V.A. Yudin, Covering a sphere and extremal properties of orthogonal polynomials, Discrete Math. Appl. 5 (1995) 371–379.