# How Good can Polynomial Interpolation on the Sphere be?

Robert S. Womersley* Ian H Sloan*

April 21, 2001

**Abstract**

This paper explores the quality of polynomial interpolation approximations over the sphere $S^{r-1} \subset \mathbb{R}^r$ in the uniform norm, principally for $r = 3$. Reimer [17] has shown there exist fundamental systems for which the norm $\|\Lambda_n\|$ of the interpolation operator $\Lambda_n$, considered as a map from $C(S^{r-1})$ to $C(S^{r-1})$, is bounded by $d_n$, where $d_n$ is the dimension of the space of all spherical polynomials of degree at most $n$. Another bound is $d_n^{1/2}(\lambda_{\text{avg}}/\lambda_{\text{min}})^{1/2}$, where $\lambda_{\text{avg}}$ and $\lambda_{\text{min}}$ are the average and minimum eigenvalues of a matrix $G$ determined by the fundamental system of interpolation points. For $r = 3$ these bounds are $(n+1)^2$ and $(n+1)(\lambda_{\text{avg}}/\lambda_{\text{min}})^{1/2}$ respectively. In a different direction, recent work by Sloan and Womersley [24] has shown that for $r = 3$ and under a mild regularity assumption, the norm of the hyperinterpolation operator (which needs more point values than interpolation) is bounded by $O(n^{1/2})$, which is optimal among all linear projections. How much can the gap between interpolation and hyperinterpolation be closed?

For interpolation the quality of the polynomial interpolant is critically dependent on the choice of interpolation points. Empirical evidence in this paper suggests that for points obtained by maximizing $\lambda_{\text{min}}$, the growth in $\|\Lambda_n\|$ is approximately $n + 1$ for $n < 30$. This choice of points also has the effect of reducing the condition number of the linear system to be solved for the interpolation weights. Choosing the points to minimize the norm directly produces fundamental systems for which the norm grows approximately as $0.7n + 1.8$ for $n < 30$. On the other hand 'minimum energy points', obtained by minimizing the potential energy of a set of $(n + 1)^2$ points on $S^2$, turn out empirically to be very bad as interpolation points.

This paper also presents numerical results on uniform errors for approximating a set of test functions, by both interpolation and hyperinterpolation, as well as by non-polynomial interpolation with certain global basis functions.

**Keywords:** Interpolation on the sphere, uniform norm, spherical polynomials, distribution of points on the sphere, Lebesgue constant.

*School of Mathematics, University of New South Wales, Sydney 2052, Australia. R.Womersley@unsw.edu.au, I.Sloan@unsw.edu.au

# 1 Introduction

In this paper we consider the construction of linear polynomial approximations on the unit sphere $S^{r-1} \subseteq \mathbb{R}^r$ from the space $\mathbb{P}_n^{(r)}$ of all spherical polynomials of degree at most $n$ (i.e. the space of all polynomials in $r$ variables restricted to $S^{r-1}$).

All the approximations to be considered are of the form $T_n f$, where $T_n$ is a projection onto $\mathbb{P}_n^{(r)}$, i.e. $T_n$ is linear and $T_n^2 = T_n$. For any linear projection $T_n$ onto $\mathbb{P}_n^{(r)}$ we have

$$\|T_n f - f\|_\infty = \|T_n(f-p) - (f-p)\|_\infty$$

for $p$ an arbitrary polynomial in $\mathbb{P}_n^{(r)}$. From this it follows immediately that

$$\|T_n f - f\|_\infty \leq (1 + \|T_n\|) E_n(f), \tag{1.1}$$

where

$$\|T_n\| = \sup_{f \in C, f \neq 0} \frac{\|T_n f\|_\infty}{\|f\|_\infty},$$

and $E_n(f)$ is the error of best uniform approximation,

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

We shall be interested in the size of $\|T_n\|$, because of its role in the upper bound (1.1).

An important example of a projective approximation is the polynomial interpolant $\Lambda_n f$, which coincides with a given continuous function $f$ at a prescribed set of points $\{x_1, \ldots, x_{d_n}\} \subseteq S^{r-1}$, where $d_n \equiv d_n^{(r)}$ is the dimension of the space of spherical polynomials of degree at most $n$. A set of points $\{x_1, \ldots, x_{d_n}\} \subseteq S^{r-1}$ is a possible set of interpolation points for the space $\mathbb{P}_n^{(r)}$ if and only if it is a 'fundamental system', that is the zero polynomial is the only member of $\mathbb{P}_n^{(r)}$ that vanishes at each point $x_j$, $j = 1, \ldots, d_n$. Now let $b_i \in \mathbb{P}_n^{(r)}$ for $i = 1, \ldots, d_n$ be a basis for $\mathbb{P}_n^{(r)}$, and define the vector valued function $\mathbf{b} : S^{r-1} \to \mathbb{R}^{d_n}$ by $\mathbf{b}(x) = [b_1(x) \cdots b_{d_n}(x)]^T$ and the interpolation matrix $B$ by

$$B = [\mathbf{b}(x_1) \cdots \mathbf{b}(x_{d_n})] \quad \text{i.e. } B_{ij} = b_i(x_j) \quad i, j = 1, \ldots, d_n. \tag{1.2}$$

Then $B$ is nonsingular if and only if the set of points $\{x_1, \ldots, x_{d_n}\}$ is a fundamental system. Given a vector $\mathbf{f} = [f(x_1), \cdots, f(x_{d_n})]^T \in \mathbb{R}^{d_n}$ of function values, the interpolation weights $v_1, \ldots, v_{d_n}$, i.e the coefficients in $\Lambda_n f = \sum_{i=1}^{d_n} v_i b_i$, are determined by the linear system

$$B^T \mathbf{v} = \mathbf{f}, \tag{1.3}$$

where $\mathbf{v} = [v_1, \ldots, v_{d_n}]^T$, and the interpolant is

$$(\Lambda_n f)(x) = \mathbf{v}^T \mathbf{b}(x) = \mathbf{f}^T B^{-1} \mathbf{b}(x). \tag{1.4}$$

The norm of the interpolation operator as a map from $C(S^{r-1})$ to $C(S^{r-1})$ (often called the *Lebesgue constant*) is

$$\begin{aligned} \|\Lambda_n\| &= \max_{\|f\|_\infty \le 1} \max_{x \in S^{r-1}} |(\Lambda_n f)(x)| && (1.5) \\ &= \max_{\|\mathbf{f}\|_\infty \le 1} \max_{x \in S^{r-1}} |\mathbf{f}^T B^{-1} \mathbf{b}(x)| && (1.6) \\ &= \max_{x \in S^{r-1}} \|B^{-1}\mathbf{b}(x)\|_1, && (1.7) \end{aligned}$$

the last step following from the existence of a continuous function $f : S^{r-1} \to \mathbb{R}$ such that $\|\mathbf{f}\|_\infty \le 1$ and $\mathbf{f} = \operatorname{sign}(B^{-1}\mathbf{b}(\bar{x}))$, where $\bar{x} \in S^{r-1}$ is a point which achieves the maximum in (1.5).

Reimer [17] and others have extensively studied the polynomial interpolation norm, yet the problem of finding a set of interpolation points that yields a good uniform approximation, or of understanding how good such approximations can be, remains elusive. One known result, from the work of Reimer [17], is that there exists a fundamental system for which $\|\Lambda_n\| \le d_n$. This property holds if the fundamental system is 'extremal', i.e. the fundamental system maximizes $|\det(B)|$, (see Section 6). This bound, which for $r = 3$ gives $\|\Lambda_n\| \le (n+1)^2$, is, as we shall see, very pessimistic. For any fundamental system $X = \{x_1, \ldots, x_{d_n}\}$ another bound on $\|\Lambda_n\|$ given by Reimer [17] has, for $r = 3$, the form $(n+1)(\lambda_{\text{avg}}/\lambda_{\text{min}})^{1/2}$ (see Section 5) where $\lambda_{\text{avg}}$ and $\lambda_{\text{min}}$ are the average and minimum eigenvalues of a certain positive-definite matrix $G$ determined by the fundamental system $X$. The ratio $\lambda_{\text{avg}}/\lambda_{\text{min}}$ depends on the choice of points $\{x_1, \ldots, x_{d_n}\}$, but beyond the fact that $\lambda_{\text{avg}}/\lambda_{\text{min}} \ge 1$ and the less obvious fact (shown by Reimer [17]) that $\lambda_{\text{avg}}/\lambda_{\text{min}} > 1$ for $r \ge 3$ and $n \ge 3$, little seems to be known about its possible dependence on $r$ and $n$. Bounds on the polynomial interpolation norm $\|\Lambda_n\|$ are considered in Section 5.

A useful comparison can be made with the 'hyperinterpolation' approximation introduced by Sloan [22]. Recently Sloan and Womersley [24] proved that for $r = 3$, and under a quadrature regularity assumption, the hyperinterpolation approximation can achieve the optimal order of growth for its operator norm among all linear projections considered as maps from $C(S^2)$ to $C(S^2)$, namely $O(n^{1/2})$. (This is the optimal rate of growth because it is achieved by the $L^2$-orthogonal projection onto $\mathbb{P}_n^{(r)}$, which is known to have the best possible uniform norm among all projections $T_n$; for details see [24]). Subsequently Reimer [20], and independently Le Gia and Sloan [11], proved the analogous result for arbitrary $r$. The hyperinterpolation approximation $L_n f$ (see Section 4.2 for more detail) is an approximation obtained from the partial sum of the Laplace (or Fourier) series for $f$, when the exact integrals in the $L^2$ inner products are approximated by a quadrature rule which has positive weights and which is exact when applied to any polynomial of degree less than or equal to $2n$. Examples of suitable quadrature rules are given in [24, 23].

An important open question is how close the rate of growth of the norm of the interpolation operator can be made, by a judicious choice of interpolation points, to the optimal order achieved by orthogonal projection and hyperinterpolation. In this paper we shall explore this question empirically for $r = 3$.

Fliege and Maier [5] report numerical integration results based on the use of interpolation points chosen to minimize the potential energy

$$\psi(x_1, \ldots, x_{d_n}) = \begin{cases} \displaystyle\sum_{i=1}^{d_n} \sum_{j=i+1}^{d_n} \log \frac{1}{\|x_i - x_j\|_2}, & r = 2 \\ \displaystyle\sum_{i=1}^{d_n} \sum_{j=i+1}^{d_n} \frac{1}{\|x_i - x_j\|_2^{r-2}}, & r \geq 3. \end{cases} \tag{1.8}$$

For the purpose of polynomial interpolation, especially in the $C \to C$ setting, these points turn out to be quite poor, as illustrated in Sections 6 and 7.

We shall see that better points may be obtained, all be it after considerable computational effort, by optimizing quantities directly associated with bounds on the norm of the interpolation operator, or even the interpolation norm itself. Explicitly, we consider fundamental systems obtained by maximizing $\lambda_{\min}$, by maximizing the determinant of the matrix $G$, and by minimizing $\|\Lambda_n\|$. The points obtained by maximizing $\lambda_{\min}$ or minimizing $\|\Lambda_n\|$ have slightly higher potential energy than those of Fliege and Maier, but several orders of magnitude larger values for $\lambda_{\min}$ and much smaller values of $\|\Lambda_n\|$. Section 6 discusses the range of strategies for choosing a good set of interpolation points on the sphere, and presents corresponding numerical results. A plausible conjecture, based on the numerical evidence, is that for $r = 3$ there exist fundamental systems for which $\|\Lambda_n\| \leq c_1 n + c_0$ where $c_1 < 1$ and $c_0 < 2$. This may be contrasted with the result $\|\Lambda_n\| \leq (n+1)^2$ of Reimer mentioned above.

Ganesh, Graham and Sivaloganathan [8] (see Section 3.1) have shown that by including also some non-polynomial basis functions they can obtain a bound of the form $C(\log n)^2$, which is asymptotically much better than the hyperinterpolation and orthogonal projection results of $O(n^{1/2})$. (However, the practical significance of this is limited, in that the asymptotic order for orthogonal projection, namely $2\sqrt{2/\pi}\, n^{1/2}$, is smaller than $(\log n)^2$ for $n \in [9, 760]$.) Jetter, Stöckler and Ward [10] have proposed non-polynomial interpolation at scattered points based on strictly positive definite radial basis functions (see Section 3.2), and establish convergence results expressed in terms of the mesh norm

$$h(X) = \sup_{x \in S^{r-1}} \operatorname{dist}(x, X), \tag{1.9}$$

where for $X = \{x_1, \ldots, x_{d_n}\}$

$$\operatorname{dist}(x, X) = \min_{j=1,\ldots,d_n} \operatorname{dist}(x, x_j) \quad \text{and} \quad \operatorname{dist}(x, y) = \cos^{-1}(x \cdot y). \tag{1.10}$$

These non-polynomial approximations provide the basis for further comparisons in Sections 5 and 7.

The main focus of this paper is on the numerical results obtained by computational optimization, rather than on the optimization techniques themselves. Techniques for finding a local solution to the eigenvalue optimization problem, which is nonsmooth when the

smallest eigenvalues become multiple, are discussed in [12, 14, 15]. Algorithms for choosing the fundamental system by minimizing $\|\Lambda_n\|$, which is a continuous minimax problem, can be found in [15] for example, with details relevant to this application in [27]. It is important to recognize that none of the optimization techniques guarantee that the final solution is a true global extremum.

Finally, because our ultimate interest is in the approximation of functions define over the sphere, the uniform norm of the error in the interpolation approximation is estimated for a number of test functions. These results dramatically illustrate the need for good choices of the points in the fundamental system for interpolation, and also show the ease with which hyperinterpolation approximations can be calculated for large $n$. Comparison with non-polynomial (but global) interpolation schemes are also made.

# 2 Polynomial interpolation

For given $n \geq 0$, let $\mathbb{P}_n^{(r)}$ be the set of spherical polynomials of degree $\leq n$ in $r$ dimensions. This section summarizes formulas for the calculation of the interpolant using three different bases: spherical harmonics, reproducing kernel zonal functions and Lagrange functions. The results in this section are not new, but are needed to discuss the relations between the bases in Section 5.

## 2.1 Spherical harmonic basis

A popular basis for $\mathbb{P}_n^{(r)}$ is the set of spherical harmonics [13]

$$\{Y_{\ell,k}^{(r)} : 1 \leq k \leq N(r,\ell), \quad 0 \leq \ell \leq n\},$$

where

$$N(r,0) = 1, \quad N(r,\ell) = \frac{2\ell + r - 2}{\ell} \binom{\ell + r - 3}{\ell - 1} \quad \text{for} \quad \ell \geq 1.$$

We shall assume that the spherical harmonics are normalized so that

$$\left(Y_{\ell,k}^{(r)}, Y_{\ell',k'}^{(r)}\right) := \int_{S^{r-1}} Y_{\ell,k}^{(r)}(x) Y_{\ell',k'}^{(r)}(x) dx = \delta_{\ell\ell'}\delta_{kk'},$$

where $dx$ denotes the surface measure on $S^{r-1}$, and $\delta_{ij}$ is the Kronecker delta. The dimension of the space $\mathbb{P}_n^{(r)}$ is

$$d_n \equiv d_n^{(r)} = \sum_{\ell=0}^{n} N(r,\ell) = N(r+1,n). \tag{2.1}$$

The addition theorem of spherical harmonics [13], which plays a key role, is

$$\sum_{k=1}^{N(r,\ell)} Y_{\ell,k}^{(r)}(x) Y_{\ell,k}^{(r)}(y) = \frac{N(r,\ell)}{|S^{r-1}|} P_\ell^{(r)}(x \cdot y), \tag{2.2}$$

where $x \cdot y$ is the inner product in $\mathbb{R}^r$, $|S^{r-1}|$ is the surface area of the unit sphere,

$$|S^{r-1}| = \frac{2\pi^{\frac{r}{2}}}{\Gamma\left(\frac{r}{2}\right)},$$

and $P_\ell^{(r)}$ is the Legendre polynomial of degree $\ell$ in $r$ dimensions, normalized by $P_\ell^{(r)}(1) = 1$. For the important special case $r = 3$ we have $N(3, \ell) = 2\ell+1$, $d_n = (n+1)^2$ and $|S^2| = 4\pi$.

For given dimension $r$ and degree $n$, let $X \equiv X_n^{(r)} = \{x_1, \ldots, x_{d_n}\} \subseteq S^{r-1}$ be a fundamental system of points on the sphere. Given an arbitrary $f \in C(S^{r-1})$, we denote by $\Lambda_n f$ the unique polynomial in $\mathbb{P}_n^{(r)}$ that interpolates $f$ at each point of the fundamental system, that is

$$\Lambda_n f \in \mathbb{P}_n^{(r)}, \quad \Lambda_n f(x_j) = f(x_j), \quad j = 1, \ldots, d_n. \tag{2.3}$$

Let $\mathbf{y} : S^{r-1} \to \mathbb{R}$ denote the vector valued function

$$\mathbf{y}(x) = \left[Y_{0,1}^{(r)}(x) \cdots Y_{n,N(r,l)}^{(r)}(x)\right]^T,$$

and define the interpolation matrix

$$A = \left[\mathbf{y}(x_1) \cdots \mathbf{y}(x_{d_n})\right], \tag{2.4}$$

i.e. $A_{N(r+1,\ell-1)+k,j} = Y_{\ell,k}^{(r)}(x_j)$ for $\ell = 0, \ldots, n$ and $k = 1, \ldots, N(r, \ell)$, with $N(r+1, -1) = 0$. The interpolant $\Lambda_n f$ can be represented as a linear combination of spherical harmonics,

$$\Lambda_n f(x) = \mathbf{v}^T \mathbf{y}(x) = \sum_{\ell=0}^{n} \sum_{k=1}^{N(r,\ell)} v_{\ell,k} Y_{\ell,k}^{(r)}(x). \tag{2.5}$$

The interpolation property (2.3) means that the coefficients must satisfy the linear system

$$A^T \mathbf{v} = \mathbf{f} \tag{2.6}$$

where $\mathbf{f} = [f(x_1) \cdots f(x_{d_n})]^T \in \mathbb{R}^{d_n}$. The matrix $A$ is nonsingular because of the assumption that $\{x_1, \ldots, x_{d_n}\}$ is a fundamental system. The matrix elements $Y_{\ell,k}^{(r)}(x_j)$ with fixed $x_j$ can be computed relatively efficiently by exploiting recurrence relations of the spherical harmonics. Consequently the time for computing $\Lambda_n f$ will often be dominated by the time needed to solve the dense linear system (2.6).

## 2.2 Reproducing kernel basis

To define the reproducing kernel basis we need first to define the reproducing kernel,

$$G_n(x, y) = \sum_{\ell=0}^{n} \sum_{k=1}^{N(r,\ell)} Y_{\ell,k}^{(r)}(x) Y_{\ell,k}^{(r)}(y), \quad x, y \in S^{r-1}, \tag{2.7}$$

the reproducing kernel property being expressed by the easily verified identity

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

Given the fundamental system $\{x_1, \dots, x_{d_n}\}$, we now define

$$g_j(x) = G_n(x_j, x), \quad j = 1, \dots, d_n, \quad x \in S^{r-1}, \tag{2.8}$$

or equivalently

$$\mathbf{g}(x) = A^T \mathbf{y}(x). \tag{2.9}$$

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

$$(g_i, g_j) = (G_n(x_i, x), G_n(x_j, x)) = G_n(x_i, x_j) = (A^T A)_{ij} \tag{2.10}$$

is nonsingular, since $A$ is nonsingular. Thus $\{g_1, \dots, g_{d_n}\}$ is a basis for $\mathbb{P}_n^{(r)}$. We shall call $g_j$ the 'kernel polynomial' associated with $x_j$.

If we represent the interpolant $\Lambda_n f$ for the fundamental system $\{x_1, \dots, x_{d_n}\}$ by

$$(\Lambda_n f)(x) = \sum_{j=1}^{d_n} w_j g_j(x) = \mathbf{w}^T \mathbf{g}(x), \tag{2.11}$$

then the weights $w_1, \dots, w_{d_n}$ are given by

$$G\mathbf{w} = \mathbf{f}, \tag{2.12}$$

where

$$G = A^T A, \tag{2.13}$$

or equivalently

$$G_{ij} = G_n(x_i, x_j), \quad i, j = 1, \dots, d_n. \tag{2.14}$$

The addition theorem for spherical harmonics (2.2) shows that that $G_n(x, y)$ is 'bizonal'

$$G_n(x, y) = \widetilde{G}_n(x \cdot y), \qquad x, y \in S^{r-1}, \tag{2.15}$$

where

$$\widetilde{G}_n(z) = \frac{1}{|S^{r-1}|} \sum_{\ell=0}^{n} N(r, \ell) P_\ell^{(r)}(z), \quad z \in [-1, 1]. \tag{2.16}$$

Thus the value of $G_n(x, y)$ depends only on the inner product $x \cdot y$ of the unit vectors $x$ and $y$, and the kernel polynomial $g_j(x)$ depends on $x$ only through $x \cdot x_j$.

The value of $\widetilde{G}_n(1)$ is, as pointed out by Reimer [17],

$$G_n(x,x) = \widetilde{G}_n(1) = \frac{d_n}{|S^{r-1}|} \quad \forall\, x \in S^{r-1}. \tag{2.17}$$

Thus the matrix $G$ in (2.14) has equal diagonal elements,

$$G_{ii} = \frac{d_n}{|S^{r-1}|} \quad i = 1, \ldots, d_n.$$

In particular, for $r = 3$ we have

$$\widetilde{G}_n(z) = \frac{1}{4\pi} \sum_{\ell=0}^{n} (2\ell+1) P_\ell(z), \tag{2.18}$$

where $P_\ell(\,\cdot\,)$ is the usual Legendre polynomial. This $r = 3$ result can be written in closed form, as pointed out by [7], in terms of the Jacobi polynomial $P_n^{(1,0)}(z)$ (in the notation of Szegö [25]) appropriate to the weight function $(1-z)$. The closed form $r = 3$ result (using [25], Equation 4.5.3) is

$$\widetilde{G}_n(z) = \frac{n+1}{4\pi} P_n^{(1,0)}(z). \tag{2.19}$$

In particular, for $r = 3$ the value at $z = 1$ is

$$\widetilde{G}_n(1) = \frac{(n+1)^2}{4\pi}.$$

For any $r \geq 3$ and any fundamental system $\{x_1, \ldots, x_{d_n}\}$ $G$ can easily be calculated by

$$G_{ij} = \widetilde{G}_n(x_i \cdot x_j) = \frac{1}{|S^{r-1}|} \sum_{\ell=0}^{n} N(r,\ell) P_\ell^{(r)}(x_i \cdot x_j). \tag{2.20}$$

The $G_{ij}$ can be efficiently evaluated using the recurrence relation for the Legendre polynomials in conjunction with (2.20), or the recurrence for the Jacobi polynomial in conjunction with (2.19).

The following simple lemma will be useful to us below.

**Lemma 2.1** *For any $x, y \in S^{r-1}$*

$$G_n(x,y) = \mathbf{g}(x)^T G^{-1} \mathbf{g}(y), \tag{2.21}$$

*where $\mathbf{g} : S^{r-1} \to \mathbb{R}^{d_n}$ is given by (2.9).*

**Proof:** For any fixed $x \in S^{r-1}$, $G_n(x,\cdot) \in \mathbb{P}_n^{(r)}$, so it can be expressed as a linear combination of the basis functions $g_i, i = 1, \ldots, d_n$, that is

$$G_n(x,\cdot) = \sum_{i=1}^{d_n} a_i(x) g_i(\cdot) = \mathbf{a}(x)^T \mathbf{g}(\cdot).$$

In particular, the two sides agree at the node points $x_j, j = 1, \ldots, d_n$, so the coefficients $a_i$ satisfy the equations

$$g_j(x) = G_n(x, x_j) = \sum_{i=1}^{d_n} a_i(x) G_{ij},$$

that is $\mathbf{g}(x) = G\mathbf{a}(x)$, or $\mathbf{a}(x) = G^{-1}\mathbf{g}(x)$. Hence

$$G_n(x,\cdot) = \mathbf{g}(x)^T G^{-1} \mathbf{g}(\cdot),$$

as required. □

The computation of the interpolant $\Lambda_n f$ via (2.11), (2.12), (2.8) and (2.20) is easy to implement, requiring only the repeated evaluation of the polynomial $\widetilde{G}_n$ and the solution of a linear system.

## 2.3 Lagrange basis

Given a fundamental system $\{x_1, \ldots, x_{d_n}\}$ of points on the unit sphere, the Lagrange polynomials $\{\ell_1, \ldots, \ell_{d_n}\}$ are defined, as usual, by

$$\ell_j \in \mathbb{P}_n^{(r)}, \quad \ell_j(x_i) = \delta_{ji}, \quad i, j = 1, \ldots, d_n. \tag{2.22}$$

For any basis $\{b_1, \ldots, b_{d_n}\}$ for $\mathbb{P}_n^{(r)}$, an explicit representation for $\ell_j$ is

$$\ell_j(x) = \frac{\det B(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_{d_n})}{\det B(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_{d_n})}, \quad j = 1, \ldots, d_n, \quad x \in S^{r-1}, \tag{2.23}$$

where $B(x_1, \ldots, x_{d_n})$ is defined by (1.2).

For given $f \in C(S^{r-1})$ the classical expression for $\Lambda_n f$ is then

$$\Lambda_n f = \sum_{j=1}^{d_n} f(x_j) \ell_j, \tag{2.24}$$

which satisfies the interpolatory property (2.3). From this it follows easily that

$$\|\Lambda_n\| = \max_{x \in S^{r-1}} \sum_{j=1}^{d_n} |\ell_j(x)|, \tag{2.25}$$

the usual form of the Lebesgue constant for interpolation. Define the vector valued function $\mathbf{l} : S^{r-1} \to \mathbb{R}^{d_n}$ by

$$\mathbf{l}(x) = \begin{bmatrix} \ell_1(x) \\ \vdots \\ \ell_{d_n(x)} \end{bmatrix}. \tag{2.26}$$

Then (2.25) may be written as

$$\|\Lambda_n\| = \max_{x \in S^{r-1}} \|\mathbf{l}(x)\|_1. \tag{2.27}$$

A concrete representation for $\mathbf{l}(x)$ is

$$\mathbf{l} = G^{-1}\mathbf{g}, \tag{2.28}$$

since the elements $\ell_j$ are then in $\mathbb{P}_n^{(r)}$, and satisfy

$$\ell_j(x_i) = \sum_{k=1}^{d_n} (G^{-1})_{jk} g_k(x_i) = \sum_{k=1}^{d_n} (G^{-1})_{jk} G_{ki} = \delta_{ji}. \tag{2.29}$$

# 3 Non-polynomial interpolation

This section summarizes two non-polynomial interpolation schemes. The first, proposed by Ganesh, Graham and Sivaloganathan [8], uses a set of basis functions which includes the spherical harmonics up to degree $n-1$ plus some non-polynomial basis functions. Remarkably, this allows them to show that the rate of growth of the interpolation operator is $O(\log^2 n)$, which is asymptotically better than the best possible rate for any projection on $\mathbb{P}_n^{(r)}$.

Another alternative, especially for scattered data, is the use of strictly positive definite kernel functions (i.e. where all coefficients in the Fourier series expansion are positive). Two examples were considered by Jetter, Stöckler and Ward [10], one of which is discussed below.

The formulae for interpolation are analogous to those in Section 2, but with $d_n$ replaced by a number $m \equiv m_n$ depending on the particular approximating space.

## 3.1 GGS interpolation

For given 'degree' $n$, Ganesh, Graham and Sivaloganathan [8] define $m = 2n^2 - 2n + 2$ basis functions

$$c_\ell^0 \quad \ell = 1, \dots, n+1, \tag{3.1a}$$
$$c_\ell^k \quad \ell = 1, \dots, n-1; \quad k = 1, \dots, n, \tag{3.1b}$$
$$s_\ell^k \quad \ell = 1, \dots, n-1; \quad k = 1, \dots, n-1, \tag{3.1c}$$

where

$$c_\ell^k(\theta,\phi) = \sin(\ell\theta)p_k(\theta)\cos(k\phi), \tag{3.2a}$$
$$s_\ell^k(\theta,\phi) = \sin(\ell\theta)p_k(\theta)\sin(k\phi), \tag{3.2b}$$

and

$$p_k(\theta) = \begin{cases} 1/\sin(\theta) & \text{if } k=0; \\ \sin(\theta) & \text{if } k \text{ is even, } k>0; \\ 1 & \text{if } k \text{ is odd.} \end{cases}$$

Note that at $\theta = 0$ (the north pole) and $\theta = \pi$ (the south pole), $c_\ell^0 = \sin(\ell\theta)/\sin(\theta)$ takes the limiting values

$$c_\ell^0(0) = \ell, \tag{3.3}$$
$$c_\ell^0(\pi) = \begin{cases} \ell & \text{if } \ell \text{ is odd;} \\ -\ell & \text{if } \ell \text{ is even.} \end{cases} \tag{3.4}$$

The basis functions (3.1) include all the spherical polynomials up to degree $n-1$ plus some non-polynomial functions.

The interpolation points are $\theta = 0$, $\theta = \pi$ and

$$\theta_i = \frac{i\pi}{n}, \quad i = 1,\ldots,n-1, \qquad \phi_j = \frac{j\pi}{n}, \quad j = 0,\ldots,2n-1. \tag{3.5}$$

For $n$ odd the maximum over the sphere $S^2$ in the expression analogous to (1.7) for the interpolation norm occurs at

$$\bar\theta = \frac{\pi}{2}, \qquad \bar\phi_j = (2j+1)\frac{\pi}{2n}, \quad j = 0,\ldots,2n-1,$$

while for even $n$ the maximum occurs at

$$\bar\theta = (n\pm 1)\frac{\pi}{2n}, \qquad \bar\phi_j = (2j+1)\frac{\pi}{2n}, \quad j = 0,\ldots,2n-1,$$

## 3.2 Interpolation using strictly positive definite functions

Jetter, Stöckler and Ward [10] analysed the use of strictly positive definite kernels for interpolation at scattered points on the unit sphere, and gave convergence results in terms of the mesh norm (1.9).

One example is a family of $C^\infty$ functions with exponentially decaying Fourier coefficients defined in terms of the generating function for the Legendre polynomials

$$\kappa(x,y) = \frac{1-z^2}{(1-2zx\cdot y+z^2)^{r/2}} = \sum_{\ell=0}^{\infty} N(r,\ell)z^\ell P_\ell^{(r)}(x\cdot y) \tag{3.6}$$

where $z \in (0,1)$ is a parameter controlling the smoothness of $\kappa$. In particular they report results using the kernel (3.6) with $z = 0.5$.

The basis functions corresponding to the point set $\{x_1, \ldots, x_m\}$ are

$$g_j(x) = \kappa(x_j, x), \quad j = 1, \ldots, m, \tag{3.7}$$

and the interpolation matrix is

$$G_{ij} = g_j(x_i) = \kappa(x_i, x_j) \quad i, j = 1, \ldots, m. \tag{3.8}$$

These strictly positive definite kernels yield a nonsingular matrix $G$ for any set of distinct points on $S^{r-1}$, but it is natural to use point sets with small values of the mesh norm, so that the points are approximately uniformly distributed over the sphere. As a specific example we use the points with spherical polar coordinates $(\theta_i, \phi_{ij})$ for $i = 0, \ldots, k$ and $j = 1, \ldots, n_i$, given by

$$\begin{array}{l} \theta_i = i\pi/k \quad i = 0, \ldots, k \\ n_0 = 1, \phi_{0,1} = 0, \quad n_k = 1, \phi_{k,1} = 0 \\ n_i = \lfloor 2\pi / \cos^{-1}\left((\cos(\pi/k) - \cos^2\theta_i)/\sin^2\theta_i\right) \rfloor \quad i = 1, \ldots, k-1 \\ \phi_{ij} = (j - 1/2)(2\pi/n_i) \quad j = 1, \ldots, n_j \quad i = 1, \ldots, k-1 \end{array} \tag{3.9}$$

which we call the FGS points (see Freeden, Gervens and Schreiner [6, Example 7.1.9]).

# 4 Non-interpolatory approximations

This section summarizes two non-interpolatory approximations: the orthogonal projection, which is known to have minimal norm amongst all linear projections onto $\mathbb{P}_n^{(r)}$, and the hyperinterpolation approximation, which was recently [24, 20, 11] shown to achieve the same optimal order.

## 4.1 Orthogonal Projection

The $L^2$ orthogonal projection of $f$ onto $\mathbb{P}_n^{(r)}$ is

$$P_n f = \sum_{\ell=0}^{n} \sum_{k=1}^{N(r,\ell)} \left(f, Y_{\ell,k}^{(r)}\right) Y_{\ell,k}^{(r)}, \tag{4.1}$$

where $(\cdot, \cdot)$ is the $L^2$ inner product on $S^{r-1}$,

$$(u, v) := \int_{S^{r-1}} u(x)v(x)dx.$$

The orthogonal projection can be written, using the addition theorem, as

$$P_n f(x) = \sum_{\ell=0}^{n} \frac{N(r,\ell)}{|S^{r-1}|} \int_{S^{r-1}} f(t) P_\ell^{(r)}(x \cdot t)dt = \int_{S^{r-1}} f(t)\widetilde{G}_n(x \cdot t)dt.$$

For $r = 3$ it follows from this that

$$\|P_n\| = 2\pi \int_{-1}^{1} |\widetilde{G}_n(z)| dz. \tag{4.2}$$

The projection $P_n$ is the minimal norm projection in the setting $C$ to $C$: that is, if $T_n$ is an arbitrary projection onto $\mathbb{P}_n^{(r)}$, then

$$\|P_n\| \le \|T_n\|.$$

This result was proved by Berman [1] for the case $r = 2$, and extended to general $r$ by Daugavet [2]; a proof for $r \ge 3$ is given by Reimer [17]. Moreover, for $r = 2$ it is known (see [3]) that $\|P_n\| \asymp \log n$, where $a \asymp b$ means that there exist constants $c_1$ and $c_2$ such that $c_1 a \le b \le c_2 a$, while for $r = 3$ (see [9])

$$\frac{\|P_n\|}{\sqrt{n}} \to 2\sqrt{\frac{2}{\pi}} \quad \text{as } n \to \infty. \tag{4.3}$$

For arbitrary $r \ge 3$, see Reimer [17, Section 11], the result is

$$\|P_n\| \asymp n^{(r-2)/2}.$$

## 4.2 Hyperinterpolation

The hyperinterpolation approximation $L_n f$ introduced in [22] is obtained by approximating the inner product in the definition (4.1) of $P_n f$ by a positive-weight quadrature rule with the property of integrating all spherical polynomials of degree at most $2n$ exactly. Thus

$$L_n f = \sum_{\ell=0}^{n} \sum_{k=1}^{N(r,\ell)} \left( f, Y_{\ell,k}^{(r)} \right)_m Y_{\ell,k}^{(r)}, \tag{4.4}$$

where $(\,\cdot\,,\,\cdot\,)_m$ is a discrete version of the inner product obtained by application of an $m$-point quadrature formula,

$$(u, v)_m := \sum_{j=1}^{m} w_j u(t_j) v(t_j),$$

and where the weights $w_j$ and points $t_j$ in the quadrature rule $Q$,

$$Qg := \sum_{j=1}^{m} w_j g(t_j) \approx \int_{S^{r-1}} g(x) dx, \tag{4.5}$$

must satisfy

$$w_j > 0, \quad t_j \in S^{r-1}, \quad j = 1, \ldots, m, \tag{4.6}$$

and

$$Qp = \int_{S^{r-1}} p(x)dx, \quad \forall\, p \in \mathbb{P}_{2n}^{(r)}. \tag{4.7}$$

From [22] it is known that $m \geq d_n$. The hyperinterpolation approximation $L_n f$ depends on the choice of the quadrature rule $Q$.

Sloan and Womersley [24] established that the norm of the hyperinterpolation operator $L_n$ in the setting $C$ to $C$ is given by (cf (4.2))

$$\|L_n\| = \max_{x \in S^{r-1}} \sum_{j=1}^{m} w_j |\widetilde{G}_n(x \cdot t_j)|, \tag{4.8}$$

and that

$$\|L_n\| \leq d_n^{1/2}.$$

For $r = 3$ this yields $\|L_n\| \leq n + 1$. Under a certain *quadrature regularity* assumption, for $r = 3$ they also show [24] the stronger result

$$\|L_n\| \asymp n^{1/2}, \tag{4.9}$$

which is optimal with respect to order. Moreover empirical evidence in [23] suggests that the constant is close to that for $P_n$.

Recently Le Gia and Sloan [11] have generalized this result to $r \geq 3$, while, independently, Reimer [20] has shown that the result holds for $r \geq 3$ without the need for any regularity assumption.

A simple expression for the hyperinterpolation approximation $L_n f$ is obtained by interchanging the order of summation in (4.4) giving

$$L_n f = \sum_{j=1}^{m} w_j f(t_j) g_j, \tag{4.10}$$

where $g_j$ denotes the kernel polynomial with axis $t_j$, that is

$$g_j(x) := G_n(x, t_j) = \widetilde{G}_n(x \cdot t_j), \quad j = 1, ..., m. \tag{4.11}$$

Equation (4.10) has a very simple structure, similar to the formula (2.11) for $\Lambda_n f$, but not requiring the solution of a linear system.

# 5 Uniform norms for interpolation

In this section we study the interpolation operator $\Lambda_n$ as a map from $C(S^{r-1})$ to $C(S^{r-1})$, and discuss various bounds on $\|\Lambda_n\|$.

Let

$$G = QDQ^T$$

be a spectral decomposition of the symmetric positive definite matrix $G$ defined by (2.14) or (2.20), with $Q$ an orthogonal matrix and $D = \mathrm{diag}(\lambda_1(G), \ldots, \lambda_{d_n}(G))$, where the eigenvalues $\lambda_j$ of $G$ are assumed to be labelled so as to satisfy

$$0 < \lambda_{\min} = \lambda_1 \leq \cdots \leq \lambda_{d_n} = \lambda_{\max}.$$

All diagonal elements $G_{ii}$ have the same value $G_n(x_i, x_i) = \widetilde{G}_n(1)$, independent of $i$, so

$$\lambda_{\mathrm{avg}} := \frac{\lambda_1 + \cdots + \lambda_{d_n}}{d_n} = \frac{\mathrm{trace}\, G}{d_n} = \widetilde{G}_n(1) = \frac{d_n}{|S^{r-1}|}. \tag{5.1}$$

The interpolation matrix $A$ in (2.4) is given by

$$A = ZD^{1/2}Q^T,$$

where $Z$ is an orthogonal matrix. Lemma 5.1 summarizes the relationships between the bases for $\mathbb{P}_n^{(r)}$ discussed in Sections 2.1, 2.2 and 2.3.

**Lemma 5.1** *For any fundamental system $\{x_1, \ldots, x_{d_n}\}$, the maps $\mathbf{y}, \mathbf{g}$, $\mathbf{l}$ and $\mathbf{q} = Q^T\mathbf{l}$ from $S^{r-1}$ to $\mathbb{R}^{d_n}$, whose components each form bases for $\mathbb{P}_n^{(r)}$, are related by*

$$\mathbf{y} = ZD^{-1/2}Q^T\mathbf{g} = A\mathbf{l} = ZD^{1/2}\mathbf{q} \tag{5.2a}$$
$$\mathbf{g} = G\mathbf{l} = QD\mathbf{q} = QD^{1/2}Z^T\mathbf{y} \tag{5.2b}$$
$$\mathbf{l} = Q\mathbf{q} = QD^{-1/2}Z^T\mathbf{y} = G^{-1}\mathbf{g} \tag{5.2c}$$
$$\mathbf{q} = Q^T\mathbf{l} = D^{-1}Q^T\mathbf{g} = D^{-1/2}Z^T\mathbf{y}. \tag{5.2d}$$

**Proof:** The properties follow from (2.9), (2.13), (2.28), the definition of $\mathbf{q} = Q^T\mathbf{l}$ and the spectral decomposition of $G$. □

The norm of the interpolation operator can now be written as

$$\begin{aligned} \|\Lambda_n\| &= \max_{x \in S^{r-1}} \|\mathbf{l}(x)\|_1 \\ &= \max_{x \in S^{r-1}} \|A^{-1}\mathbf{y}(x)\|_1 \\ &= \max_{x \in S^{r-1}} \|G^{-1}\mathbf{g}(x)\|_1 \\ &= \max_{x \in S^{r-1}} \|Q\mathbf{q}(x)\|_1. \end{aligned} \tag{5.3}$$

The value of $\|\Lambda_n\|$ depends on the fundamental system $\{x_1, \ldots, x_{d_n}\}$. One knows that $\|\Lambda_n\|$ can be made arbitrarily large if the fundamental system is badly chosen. The interesting question is how small $\|\Lambda_n\|$ can be made by a judicious choice of fundamental system. Little is known about this question.

One known result, from the work of Reimer [21, 17, 18], is that there exists a fundamental system with $\|\Lambda_n\| \le d_n$. This property holds if the fundamental system is 'extremal'. A fundamental system $X = \{x_1, \dots, x_{d_n}\}$ is extremal if it maximizes $|\det B(x_1, \dots, x_{d_n})|$, where $B(x_1, \dots, x_{d_n})$ is defined by (1.2) and $\{b_1, \dots, b_{d_n}\}$ is any basis. For an extremal fundamental system the explicit representation (2.23) gives immediately

$$\|\ell_j\|_\infty = 1 \quad j = 1, \dots, d_n,$$

and hence from (2.25)

$$\|\Lambda_n\| \le d_n.$$

This bound, which for $r = 3$ is $\|\Lambda_n\| \le (n+1)^2$, is very pessimistic.

Another bound on $\|\Lambda_n\|$ (see Corollary 2 of [16]), which is obtained by considering the 'Lagrangian square sum' $\sum_{j=1}^{d_n} \ell_j(x)^2$, is

$$\|\Lambda_n\| \le d_n^{1/2} \sqrt{\lambda_{\text{avg}}/\lambda_{\text{min}}}. \tag{5.4}$$

Reimer notes in particular (in Corollary 3 of [16]) that in the special case that the eigenvalues are all equal, i.e. $\lambda_1 = \cdots = \lambda_{d_n}$, the result reduces to

$$\|\Lambda_n\| \le d_n^{1/2}. \tag{5.5}$$

However, it is known (see [17]), that for $r \ge 3$ and $n \ge 3$ the eigenvalues can **not** be equal.

At the present time it is an open question whether the bound (5.4) can be improved. In the case $r = 3$ the bound is $(n+1)(\lambda_{\text{avg}}/\lambda_{\text{min}})^{1/2} \ge n+1$. A bound with a different character is given by the following theorem.

**Theorem 5.2** *Let $r = 3$. For each $n \ge 1$ let $\{x_1, \dots, x_{d_n}\}$ be a fundamental system with the property that*

$$|x_j - x_k| \ge c_1/\sqrt{d_n} = c_1/(n+1), \quad j \ne k,$$

*for some positive constant $c_1$ independent of $n, j$ and $k$. Then there exists $c_2 > 0$ such that*

$$\|\Lambda_n\| \le c_2 n^{1/2} \left(\frac{d_n}{4\pi}\right) \|G^{-1}\|_1 = c_2 n^{1/2} \lambda_{\text{avg}} \|G^{-1}\|_1,$$

*where $\lambda_{\text{avg}}$ given in (5.1) is the mean of the eigenvalues of the $d_n$ by $d_n$ matrix $G$ defined by (2.14).*

**Remark 1** If equal eigenvalues were possible (which of course is not the case), then $\lambda_1 = \cdots = \lambda_{d_n} = \lambda_{\text{avg}}$ would give $G = \lambda_{\text{avg}} I$ and $\|G^{-1}\|_1 = \lambda_{\text{avg}}^{-1}$, so the result of the theorem would be

$$\|\Lambda_n\| \le c_2 n^{1/2},$$

while (5.5) would give $\|\Lambda_n\| \le n+1$, which is asymptotically worse.

**Proof:** The Lebesgue constant (2.25) may be written

$$\begin{aligned}\|\Lambda_n\| &= \max_{x\in S^{r-1}} \|\mathbf{l}(x)\|_1 = \max_{x\in S^{r-1}} \|G^{-1}\mathbf{g}(x)\|_1 \\ &\le \|G^{-1}\|_1 \max_{x\in S^{r-1}} \|\mathbf{g}(x)\|_1 \\ &= \frac{d_n}{4\pi}\|G^{-1}\|_1 \max_{x\in S^{r-1}} \frac{4\pi}{d_n}\sum_{k=1}^{d_n} |g_k(x)|. \end{aligned} \tag{5.6}$$

At this stage we may appeal to the arguments used to prove Theorem 5.4 in [24], if we define a $d_n$-point quadrature rule with equal weights by

$$Q_{d_n} f := \frac{4\pi}{d_n}\sum_{k=1}^{d_n} f(x_k). \tag{5.7}$$

This rule is exact for constant functions. Moreover, by Proposition 6.1 of [24] the family of quadrature rules $\{Q_{d_n}\}$ has the quadrature regularity property. The result follows by using the estimate for $\|L_n\|$ in [24, Theorem 5.4] to estimate the last factor in (5.6). □

Numerical evidence suggests that the bound in Theorem 5.2 is in practice weaker than Reimer's eigenvalue bound (5.4).

# 6 Good fundamental systems

This section considers different criteria which may be used to select a good fundamental system. In particular, we use the labels ME, MD, EV and MN to denote the fundamental systems $X = \{x_1, \dots, x_{d_n}\}$ selected according to the criteria in Table 1.

| Label | Criteria |
|---|---|
| ME | Minimize the potential energy $\psi(X)$ defined in (1.8), as determined by Fliege and Maier [5]. |
| MD | Maximize $\log\det(G(X))$), [21, 17]. |
| EV | Maximize $\lambda_{\min}(G(X))$. |
| MN | Minimize the norm $\|\Lambda_n\| = \max_{x\in S^{r-1}} \|G^{-1}\mathbf{g}(x)\|_1$. |

Table 1: Criteria for choosing fundamental systems

Proposition 6.1 summarizes the situation if the off-diagonal elements in the symmetric positive definite matrix $G$ could be chosen freely. It is just an application of the arithmetic mean – geometric mean – harmonic mean inequality

$$\frac{1}{m}\sum_{i=1}^{m}\lambda_i \ge \left(\prod_{i=1}^{m}\lambda_i\right)^{1/m} \ge \left(\frac{1}{m}\sum_{i=1}^{m}\frac{1}{\lambda_i}\right)^{-1},$$

which holds with equality if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_m$.

**Proposition 6.1** *Let $\mathcal{G}$ denote the set of real symmetric positive definite $m$ by $m$ matrices with fixed diagonal elements $G_{ii} = \lambda_{\text{avg}} > 0$ for $i = 1, \ldots, m$. The problems*

$$\max_{G \in \mathcal{G}} \ \lambda_{\min}(G), \tag{6.1a}$$

$$\max_{G \in \mathcal{G}} \ \det(G), \tag{6.1b}$$

$$\max_{\substack{A \in \mathbb{R}^{m \times m}, \\ A^T A \in \mathcal{G}}} \ \det(A), \tag{6.1c}$$

$$\min_{G \in \mathcal{G}} \ \frac{\lambda_{\text{avg}}(G)}{\lambda_{\min}(G)}, \tag{6.1d}$$

$$\min_{G \in \mathcal{G}} \ \text{cond}(G), \tag{6.1e}$$

$$\min_{G \in \mathcal{G}} \ \|G\|_1, \tag{6.1f}$$

$$\min_{G \in \mathcal{G}} \ \text{trace}\, G^{-1}, \tag{6.1g}$$

*all have optimal solution $G = \lambda_{\text{avg}} I$.*

Proposition 6.1 also holds if $\mathcal{G}$ is the set of real symmetric positive definite matrices with fixed trace $G = m\lambda_{\text{avg}}$.

For $r = 3$ and $n = 1$ the fundamental system

$$X = [x_1 \ x_2 \ x_3 \ x_4] = \begin{bmatrix} 0 & \frac{2\sqrt{2}}{3} & -\frac{\sqrt{2}}{3} & -\frac{\sqrt{2}}{3} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ 1 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

(which consists of the vertices of a regular tetrahedron) produces $G = \lambda_{\text{avg}} I = (1/\pi) I$ and $\|\Lambda_1\| = 2$. However, as already noted, it is not possible for $r \geq 3$ and $n \geq 3$ to choose a fundamental system of interpolation points to make $G = \lambda_{\text{avg}} I$. Thus by varying the fundamental system $x_1, \ldots, x_m$ a subset of the matrices in $\mathcal{G}$ can be generated, and optimizing the various criteria (6.1) produces different results. The matrix $G$ depends only on the angles between the point of the fundamental system, so is invariant under an arbitrary rotation of the fundamental system. Thus, as in Fliege and Maier [5], the fundamental system is conveniently normalized by taking the first point at the north pole and the second point on the prime meridian. For $r = 3$, using spherical coordinates $\theta_j \in [0, \pi]$ and $\phi_j \in [0, 2\pi)$ for $j = 1, \ldots, d_n$, with $\theta_1 = 0$ (with $\phi_1 = 0$) and $\phi_2 = 0$, this gives a total of $2d_n - 3$ variables to define the fundamental system.

Swapping any two points corresponds to swapping two rows of $A$, so $\det A$ changes sign. Since $A(x_1, \ldots, x_{d_n})$ is continuous, there are points where $A$ and $G$ are singular and the norm is infinite. Thus there may be many different local optima, often separated by arbitrarily bad objective values. Guaranteeing a global optimum is a virtually intractable problem, so the EV and MD points were obtained by starting the optimization from the minimum energy points of Fliege and Maier. The MN points were obtained by starting from the EV points. Due to the size ($n = 29$ gives 1797 variables) and difficulty (both

$\max \lambda_{\min}(G)$ and $\max_{x \in S^{r-1}} \|G^{-1}\mathbf{g}(x)\|_1$ are nonsmooth functions of the fundamental system $X$) only approximate local optimizers can be found. Most difficult of all is the criterion MN.

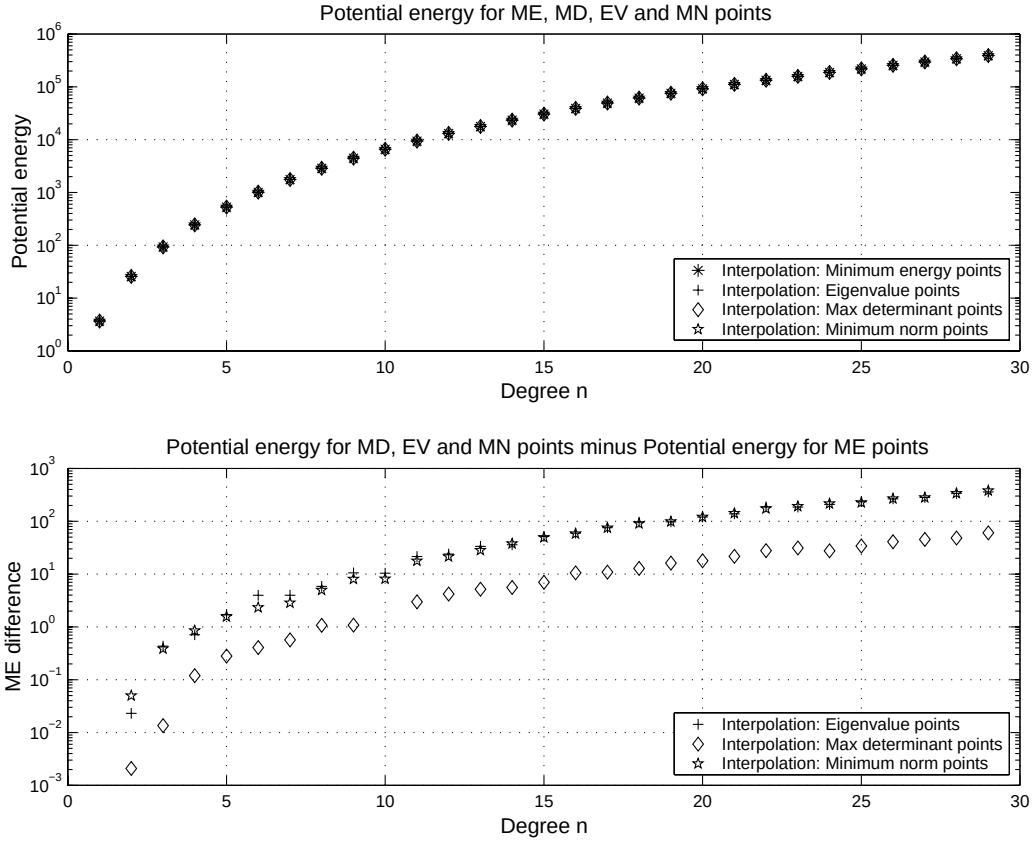


Figure 1: Potential energy for fundamental systems

Figure 1 (in its upper part) shows the potential energy (1.8) as a function of $n$ for each of the point sets ME (the starting set due to Fliege and Maier), EV, MD and MN. A striking fact seen in Figure 1 is that the potential energy for each $n$ has hardly changed from its initial value, even though (as we will see) in other aspects the new point sets MN, EV and MD all have much better properties for interpolation. Because the curves in the upper part of Figure 1 are not distinguishable to the eye, in the lower part of the figure we show (on a scale that is three orders of magnitude larger) the difference between the potential energies for the EV, MD and MN points on the one hand, and the starting ME points on the other. (For $n = 10$ the MD points have potential energy approximately 0.8 *less* than the potential energy for the ME points, so the negative difference does not appear on a log scale. This and later evidence in Figure 7 supports the conjecture that the $n = 10$ ME point set of [5] is in some way flawed.) A broad conclusion seems to be that the potential energy is a remarkably insensitive measure of the quality of a point set for

interpolation.

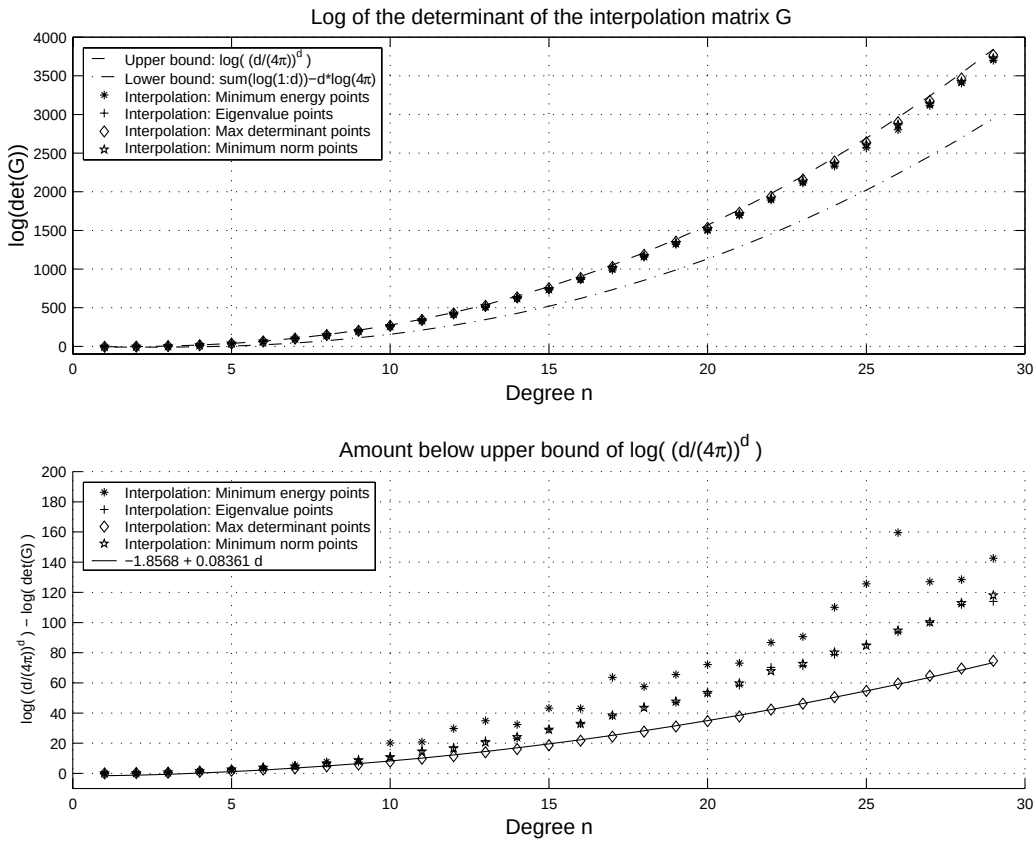


Figure 2: $\log\det G$ for fundamental systems

The extremal fundamental systems of Reimer (here labelled MD) are chosen to maximize $\det G$. Proposition 6.1 provides the upper bound

$$\det G = \prod_{i=1}^{d_n} \lambda_i(G) \leq \left(\lambda_{\text{avg}}\right)^{d_n} = \left(\frac{d_n}{|S^{r-1}|}\right)^{d_n}.$$

The right hand side grows so rapidly that it is not representable in a floating point number system for modest $n$. For instance for $r = 3$ and $n = 15$, $(d_n/4\pi)^{d_n} > 10^{335}$ which is not representable in IEEE floating point arithmetic. Thus one must work with

$$\log\det G = \sum_{i=1}^{d_n} \log\lambda_i(G) \leq d_n \log\left(\frac{d_n}{4\pi}\right) \tag{6.2}$$

for $r = 3$. Figure 2 plots $\log\det G$ and $(d_n \log\left(\frac{d_n}{4\pi}\right) - \log\det G)$, the amount below the upper bound (6.2), for the ME, MD, EV and MN points. Reimer and Sündermann [21]

gave an exchange algorithm to calculate extremal fundamental systems. Here we directly maximized $\log\det G(X)$ using a normalized spherical parametrization of the point set $X$. This gave slightly larger (and hence better) values of $\log\det G(X)$ than the fundamental systems reported in [21] for $n = 3, 4, 5, 6$.

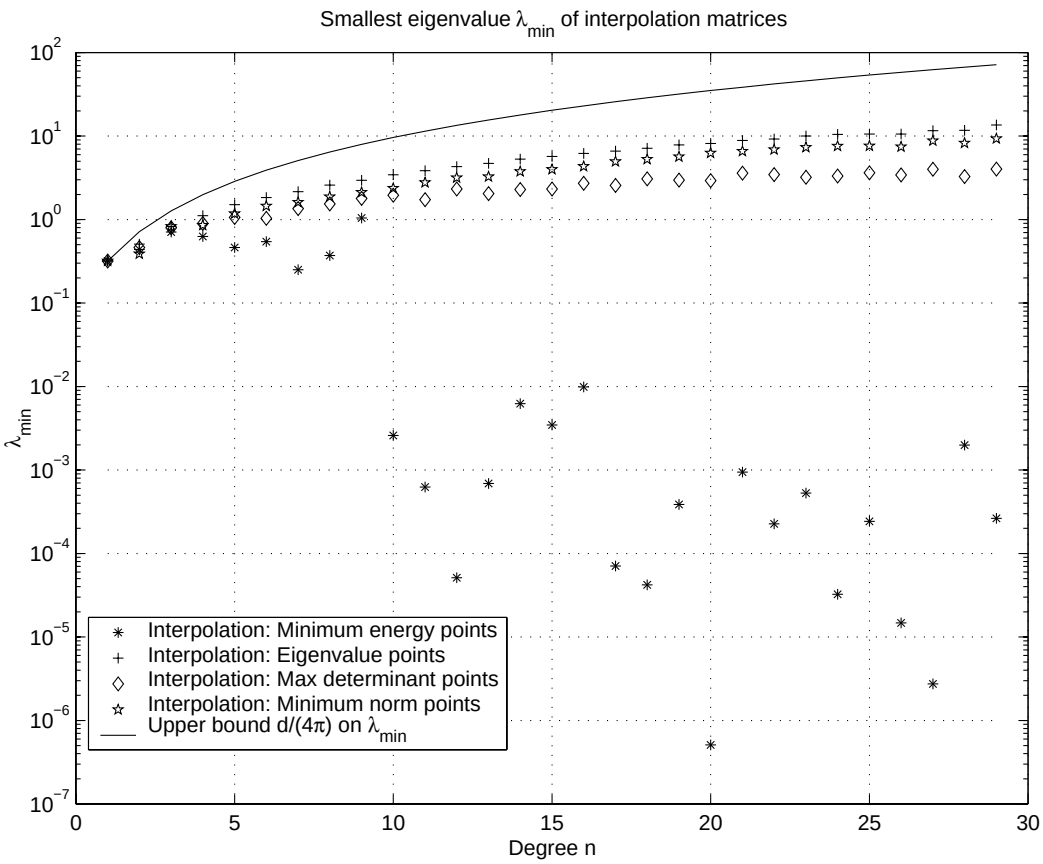


Figure 3: Smallest eigenvalues of the interpolation matrices

Reimer [19] has also shown that there exists a fundamental system $\bar{X}$ with

$$\det \bar{G} = \frac{d_n!}{|S^{r-1}|^{d_n}},$$

which for $r = 3$ gives

$$\log\det G(\bar{X}) = \sum_{i=1}^{d_n} \log(i) - d_n \log(4\pi)$$

and provides a lower bound on the maximum determinant. The numerical evidence (see Figure 2), shows that this lower bound is pessimistic, and that the maximal determinant

(i.e. MD) points are much closer to the upper bound, in that for $n < 30$

$$\log\det G(X)) \approx d_n \log d_n/(4\pi) - (a_0 + a_1 d_n),$$

where $a_0 \approx -1.9$ and $a_1 \approx 0.08$.

Reimer [19] also suggested choosing the fundamental system by solving (6.1g), for which Proposition 6.1 gives

$$\text{trace } G^{-1} = \sum_{i=1}^{d_n} \frac{1}{\lambda_i(G)} \geq |S^{r-1}|.$$

The bound (5.4), and the fact that $\lambda_{\text{avg}} = d_n/|S^{r-1}|$, suggest that an alternative is to choose the fundamental system to minimize $\lambda_{\text{avg}}/\lambda_{\text{min}}$ by solving

$$\max_{\{x_1,\dots,x_{d_n}\}} \lambda_{\text{min}}(G). \tag{6.3}$$

As the sum of the eigenvalues of $G$ is fixed,

$$\text{trace } G = \sum_{i=1}^{d_n} \lambda_i = \frac{d_n^2}{|S^{r-1}|}$$

this also has the effect of improving the spectral condition number

$$\text{cond}_2(G) = \frac{\lambda_{\text{max}}(G)}{\lambda_{\text{min}}(G)}.$$

The EV points were chosen by solving (6.3) starting from the ME points. This dramatically increases the smallest eigenvalue $\lambda_{\text{min}}(G)$, while keeping the largest eigenvalue $\lambda_{\text{max}}(G)$ approximately the same. Hence the condition number of $G$ is also dramatically reduced (see Figure 6). The smallest eigenvalue for the ME, EV, MD and MN points are plotted in Figure 3. Note that the EV points have smallest eigenvalues up to $10^7$ times larger than the minimum energy (or ME) points. The small eigenvalues occurring for the ME points adversely affect the accuracy with which the linear system for the interpolation weights can be solved, as well as the bound on the norm of the interpolation operator and the accuracy of the resulting interpolant.

A fundamental system $X = \{x_1, \dots, x_{d_n}\}$ is said to be 'optimal for interpolation' if it minimizes the norm of the interpolation operator, that is

$$X = \operatorname*{argmin}_{\{x_1,\dots,x_{d_n}\}} \|\Lambda_n\|.$$

From (2.27) this is equivalent to solving the minimax problem

$$\min_{\{x_1,\dots,x_{d_n}\}} \max_{x\in S^{r-1}} \|G^{-1}\mathbf{g}(x)\|_1. \tag{6.4}$$

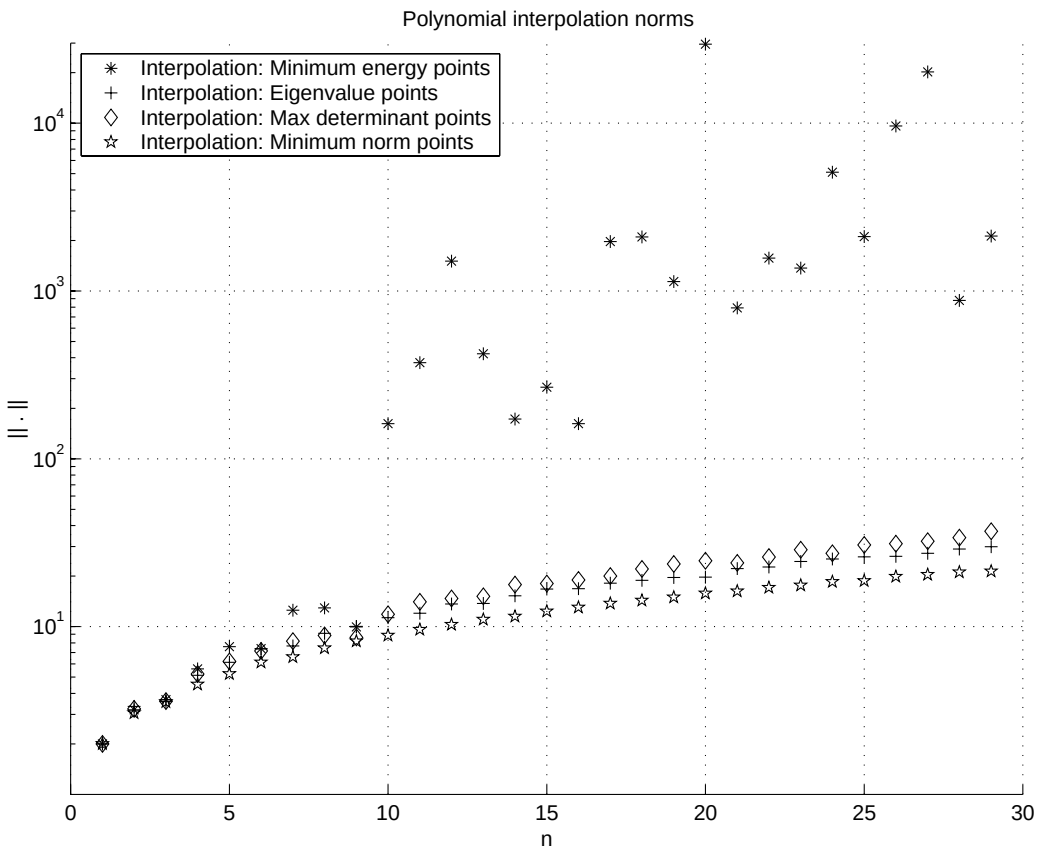


Figure 4: Norms of the polynomial interpolation operators

This is a continuous minimax problem in which the inner maximization requires the calculation of the global maximum over the sphere. Efficient methods for solving (6.4) require all local maxima close to the global maximum. Moreover as the norm is minimized many points are expected to achieve the global maximum (for instance with the MN points for $n \geq 20$ there are typically several hundred local maxima within 1% of the global maximum). See [27] for more details.

The norms $\|\Lambda_n\|$ are plotted in Figures 4 and 5 as functions of $n$. For $r = 3$,

$$\|\Lambda_n\| = \max_{x \in S^2} \|G^{-1}\mathbf{g}(x)\|_1.$$

The function $\|G^{-1}\mathbf{g}(x)\|_1$ is a continuously differentiable function of $x \in S^2$ at all points $x$ where $(G^{-1}\mathbf{g}(x))_i \neq 0$ for $i = 1, \ldots, d_n$. Thus the norm is calculated by first using a grid of uniformly spaced points on $S^2$, identifying potential local maxima within 1% of the global maximum on the grid, then accurately finding all local maxima from these starting points using a smooth unconstrained maximization routine, and taking the largest of these local maxima as $\|\Lambda_n\|$. A notable feature of the results in Figure 5 is that $\|\Lambda_n\| \approx n + 1$ for the EV points, and $\|\Lambda_n\| \approx 0.7n + 1.8$ for the MN points.

Theorem 5.2 gave a bound for $\|\Lambda_n\|$ of the form $cn^{1/2}\lambda_{\text{avg}}\|G^{-1}\|_1$. For all the point sets EV, MD, ME and EV the term $\lambda_{\text{avg}}\|G^{-1}\|_1$ grew sufficiently to make this bound worse than the bound (5.4).

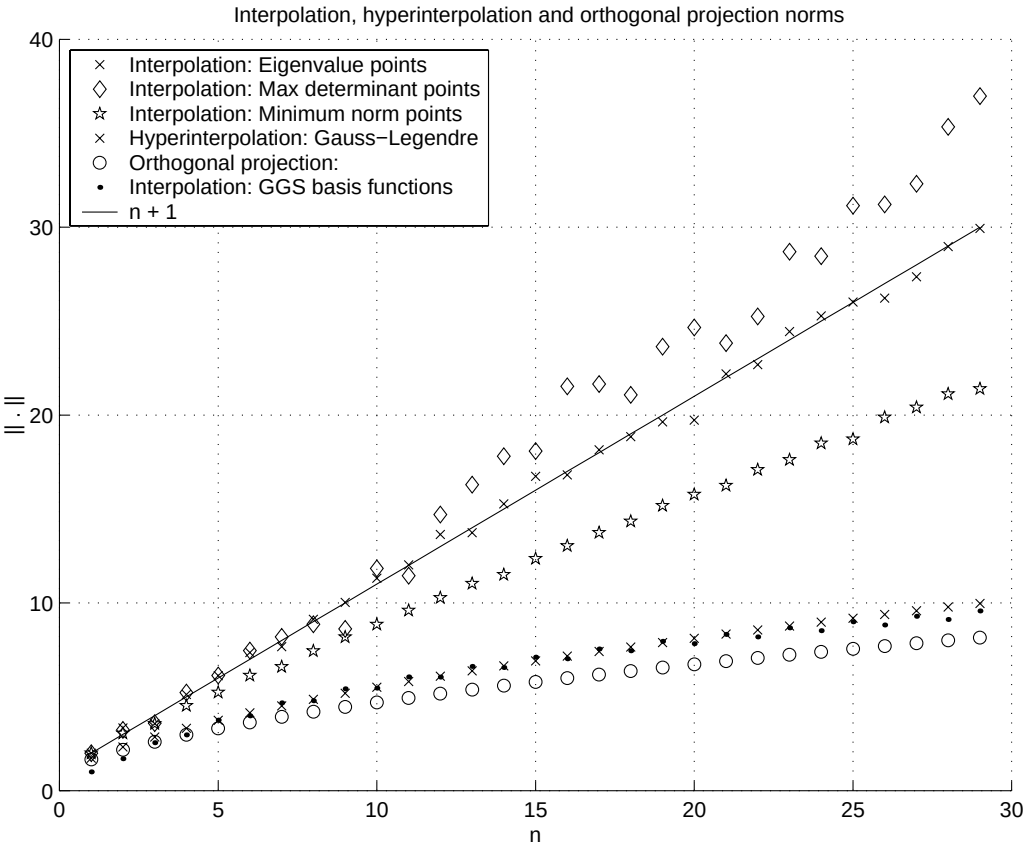


Figure 5: Norms for interpolation, hyperinterpolation and orthogonal projection

The approximately linear growth in the interpolation norms for the EV and MN points is still above the $O(n^{1/2})$ growth for the hyperinterpolation and orthogonal projection, as illustrated in Figure 5. The addition of the some non-polynomial basis functions, as in Section 3.1, also produces low interpolation norms, at least asymptotically. Figures 4 and 5 plot the norms against the degree $n$ of the polynomials. It must be born in mind that the hyperinterpolation schemes and the GGS interpolation method use 2 to 4 times as many points compared with polynomial interpolation, as shown in Table 2.

For all interpolation schemes, the condition number of the the linear system solved to get the interpolation weights affects the accuracy of the interpolant, especially for large $n$. Even though interpolation by positive definite functions (see Section 3.2) is theoretically defined for any set of distinct points, the linear system (1.3) must still be solved for the interpolation weights. For the well-distributed points (3.9) and the kernel (3.6) for small $z = 0.1$ and $z = 0.5$ the condition number of the interpolation matrix grows very rapidly,

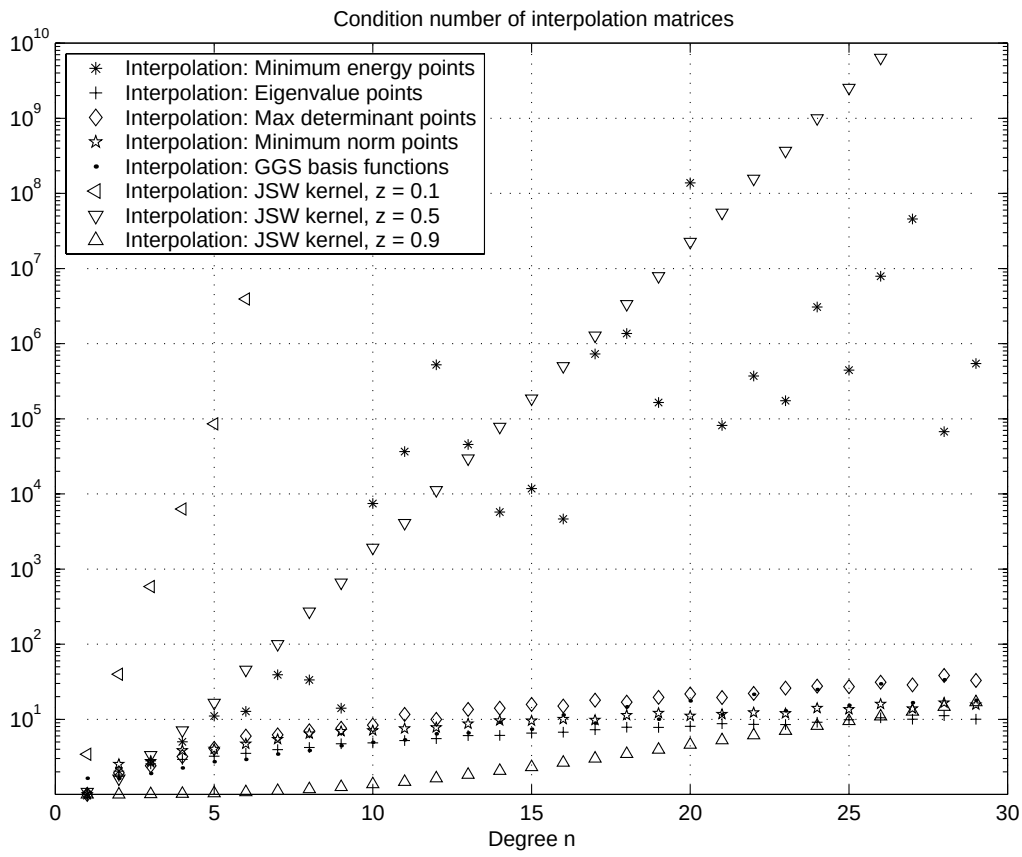


Figure 6: Condition number of interpolation matrices

making it useless for even moderate numbers of points. In contrast, the hyperinterpolation approximations (see Section 4.2) use 2–4 times the number of points (see Table 2), but do not require the solution of a linear system for the weights, provided the underlying quadrature rule (4.5) is known in advance. Another major advantage of the hyperinterpolation approximation is that the points and weights can be generated with very little work if one of the standard quadrature rules is used, while choosing good interpolation points is a significant computational task for large $n$. On the other hand, good interpolation points once found can be stored for future use.

The mesh norms (1.9) for the ME, MD, EV, MN, GGS, and FGS points are plotted in Figure 7. The behaviour of the mesh norms is similar when plotted against $n$, except for the anomalous result for the ME points for $n = 10$.

| Method | Point Set | Number of points |
|---|---|---|
| Polynomial Interpolation | Fundamental System | $(n+1)^2$ |
| GGS Interpolation | GGS points | $2n^2 - 2n + 2$ |
| JSW Interpolation | FGS points | $\approx 2 + 4n^2/\pi$ |
| Hyperinterpolation | Gauss-Legendre | $2(n+1)^2$ |
| Hyperinterpolation | Clenshaw-Curtis | $4n^2 + 2n$ |
| Hyperinterpolation | Fejér | $4n^2 + 6n + 2$ |

Table 2: Number of points as a function of $n$

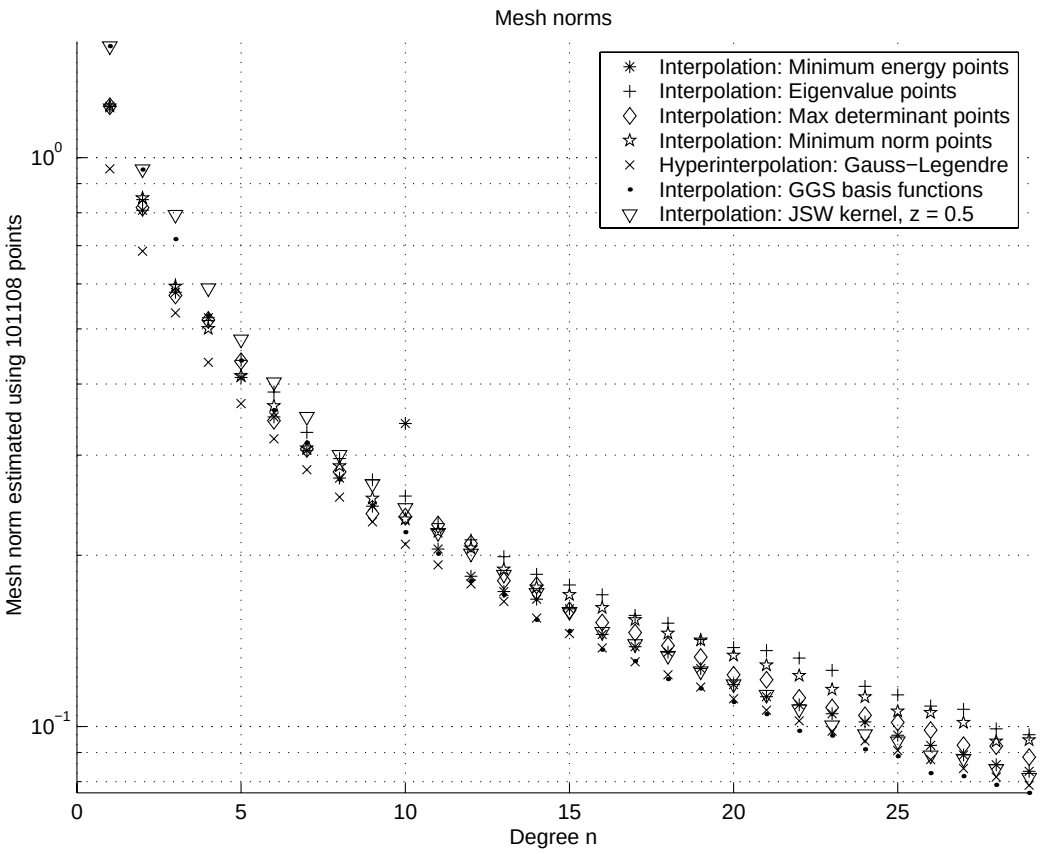


Figure 7: Mesh norms

# 7 Uniform errors on test functions

In the preceding sections we have studied the uniform norm of the various projection operators, which through (1.1) provide an upper bound on the uniform error of approximation of a continuous function $f$. However because (1.1) is a one-sided bound, arguments so far present an incomplete picture of the usefulness of the various approximation schemes in practice.

We therefore consider the uniform error for polynomial interpolation, non-polynomial interpolation and hyperinterpolation approximations on a set of test functions. The test functions used are

$$\begin{aligned}
f_1(x) &= x_1 x_2 x_3, \\
f_2(x) &= e^{x_1}, \\
f_3(x) &= \tfrac{1}{10} e^{x_1+x_2+x_3}, \\
f_4(x) &= -5\sin(1+10x_3), \\
f_5(x) &= 1/(101-100x_3), \\
f_6(x) &= \sum_{i=1}^{5} \alpha_i e^{-\beta_i \mathrm{dist}(x,y_i)^{2p_i}}, \qquad (7.1) \\
f_7(x) &= \|x\|_1/10, \\
f_8(x) &= 1/\|x\|_1, \\
f_9(x) &= \sin^2(1+\|x\|_1)/10, \\
f_{10}(x) &= \begin{cases} h_0 \cos^2(\frac{\pi}{2}\mathrm{dist}(x,x_0)/R) & \text{if } \mathrm{dist}(x,x_0) < R \\ 0 & \text{if } \mathrm{dist}(x,x_0) \geq R, \end{cases} \qquad (7.2)
\end{aligned}$$

where the geodesic distance between two points $x, y$ on the unit sphere is

$$\mathrm{dist}(x,y) = \cos^{-1}(x \cdot y),$$

and the parameters for $f_6$ are given in Table 3.

| $i$ | $y_{i,1}$ | $y_{i,2}$ | $y_{i,3}$ | $\alpha_i$ | $\beta_i$ | $p_i$ |
|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 2 | 5 | 1 |
| 2 | 0.932039 | 0 | 0.362358 | 0.5 | 7 | 1 |
| 3 | −0.362154 | 0.619228 | 0.696707 | −2 | 6 | 2 |
| 4 | 0.904035 | 0.279651 | −0.323290 | −2 | 5 | 1 |
| 5 | −0.0479317 | −0.424684 | −0.904072 | 0.2 | 2.1 | 1 |

Table 3: Parameters for test function $f_6$

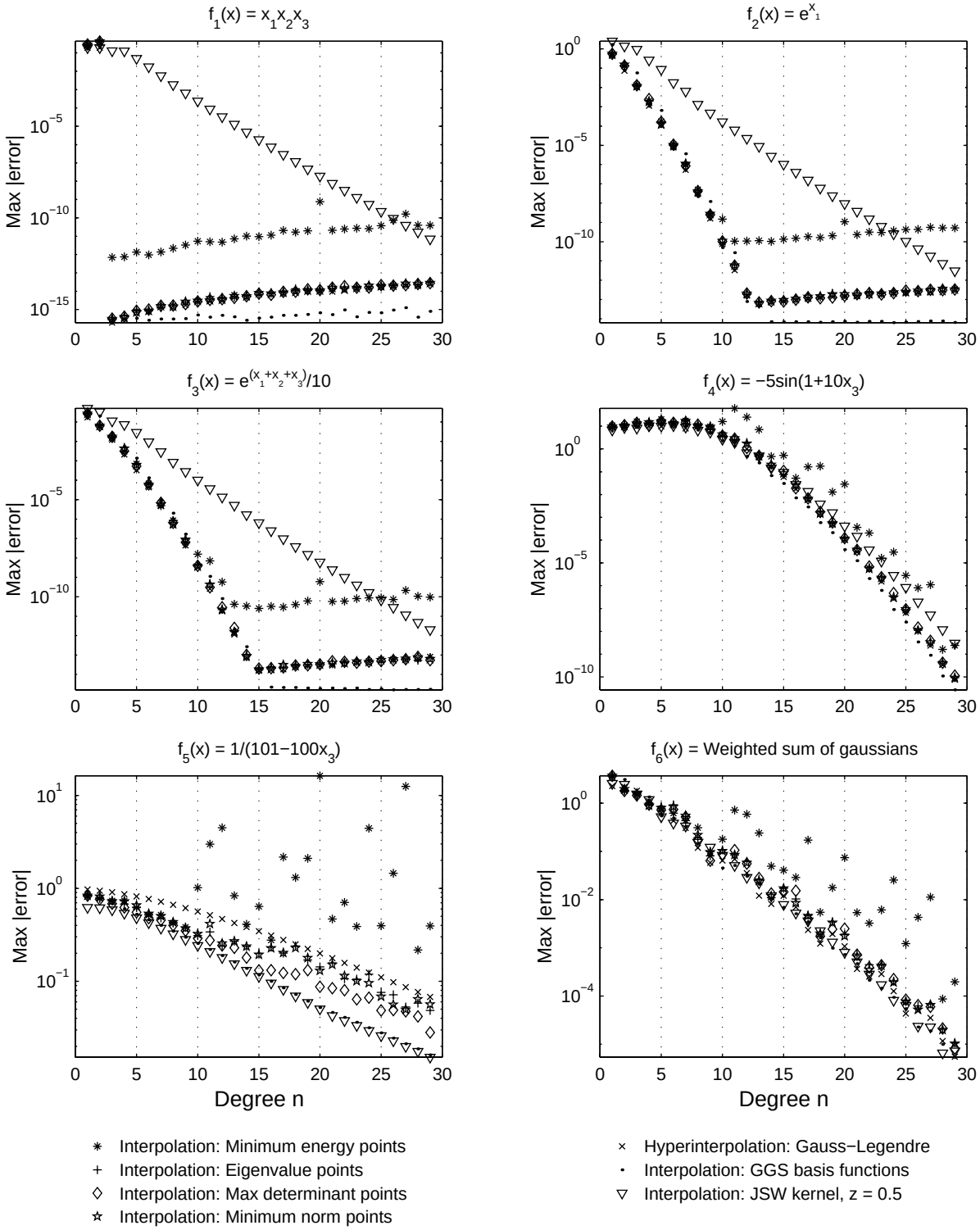


Figure 8: Uniform error for smooth test functions

The functions $f_1, \ldots, f_5$ and $f_7, f_8, f_9$ were used by Fliege and Maier [5] to test the quality of their numerical integration scheme, which is based on integration of the polynomial interpolant through their calculated points. The function $f_1$, a polynomial of low degree, is useful for code checking. Functions $f_2, \ldots, f_5$ are analytic over $S^2$ (although $f_5$ has a pole just off the surface of the sphere at $x = [0\ 0\ 1.01]$),

The function $f_6$, defined in (7.1), was introduced by Jetter, Stöckler and Ward [10]. It is a weighted sum of five Gaussian-like functions with the parameters $\alpha_i$, $\beta_i$, $p_i$ and points $y_i \in S^2$ given in Table 3. The function $f_6(x)$ is analytic, with a global maximum near $y_1$ and minima near $y_3$ and $y_4$. The term centred at $y_3$ decays radially like $e^{-\rho^4}$, and is the hardest component to approximate accurately.

The functions $f_7, f_8, f_9$ have only $C^0$ continuity (in particular they are not continuously differentiable at points where any component of $x$ is zero).

The cosine cap $f_{10}$, defined in (7.2), is part of a standard test set [26] for numerical approximations to the shallow water equations in spherical geometry. The point $x_0 \in S^2$ is the centre of the cap with radius $R$ and height $h_0$. The results in Figure 9 used a centre $x_0$ with polar coordinates $(\theta_0, \phi_0) = (\pi/4, 5\pi/4)$, radius $R = 1/3$ and amplitude $h_0 = 1$. Other parameters can easily be used. The cosine cap has local support and is only once continuously differentiable at a distance $R$ from the centre of the cap.

The uniform norm of the errors is estimated by

$$\|f - L_n f\|_\infty \approx \max_{x \in X} |f(x) - L_n f(x)|, \tag{7.3}$$

where $X$ is a large, but finite, set of well distributed points over the sphere, for instance (3.9), with 101105 points, corresponding to $\ell = 281$.

The calculated uniform errors of the interpolation approximations using the minimum energy and eigenvalues points, and the hyperinterpolation approximation using the Clenshaw-Curtis rule, are given in Figures 8 and 9. The test functions in Figure 8, being at least twice continuously differentiable, should be amenable to approximation by polynomials. The test functions in Figure 9 have discontinuities in their first derivatives, except for the cosine cap which has a discontinuity in its second derivatives.

The less than ideal behaviour of the minimum energy (or ME) points is clear, especially for the more difficult test functions in Figure 9. The EV points generally perform well, as do the MD and MN points. Also noteworthy is the good behaviour of the GGS interpolation scheme, in which the interpolation points are fully prescribed.

As Reimer [18] did for the MD points and Fliege-Maier [5] did for the ME points, any of the point sets discussed can be used for quadrature over the surface of the sphere by employing

$$\int_{S^2} f(x)\, dx \approx \mathbf{w}^T \mathbf{f}, \tag{7.4}$$

where $G\mathbf{w} = \mathbf{e}$, $\mathbf{e} \in \mathbb{R}^{d_n}$ is the vector of all ones, and $\mathbf{f} = [f(x_1), \ldots, f(x_{d_n})]^T$. Some of the numerical difficulties seen in [4, 5] when using the ME points for numerical quadrature can

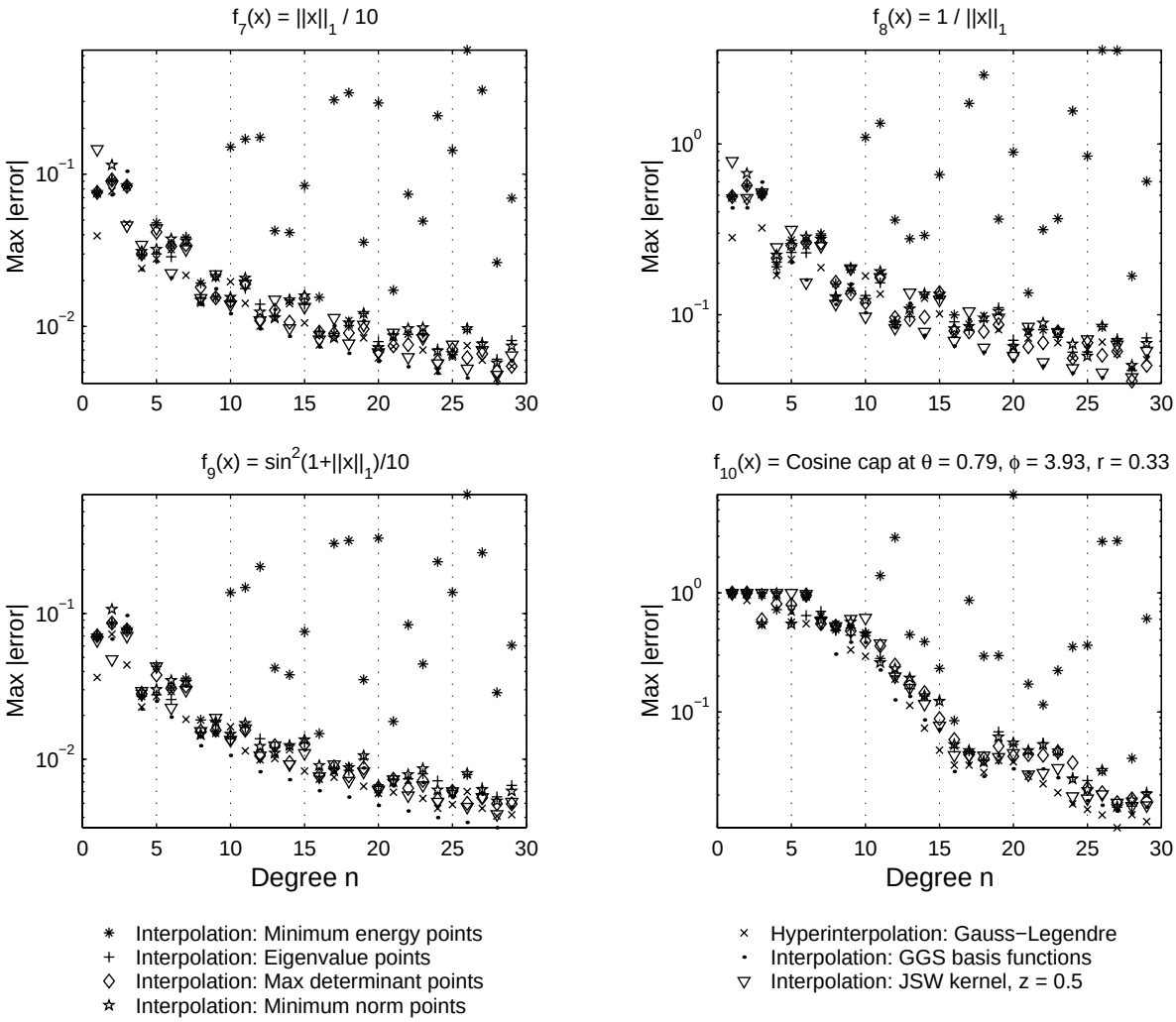


Figure 9: Uniform error for nonsmooth test functions

be attributed to the poor conditioning of the linear system $G\mathbf{w} = \mathbf{e}$ that must be solved for the weights.

The MD, EV and MN points and associated weights $\mathbf{w}$ in (7.4), along with illustrations of actual point distributions, are available from the web site
`http://www.maths.unsw.edu.au/~rsw/Sphere/`.

# 8 Conclusions

The conclusions expressed here, based largely on numerical experiments, must necessarily be cautious. In particular the points sets ME, MD, EV and MN are all approximate *local* optimizers of their respective criteria, and there is no guarantee that they are the global

optimizers.

The safest conclusion is that the choice of interpolation points really matters! In particular the minimum energy points, which are regular in appearance and (nearly) optimal from one point of view, are nevertheless remarkably unsuitable for use as interpolation points for polynomial interpolation.

Among the polynomial interpolation schemes, the eigenvalue (EV) and maximum determinant (MD) are consistently effective. So too are the minimum norm (MN) points: these are the most expensive to obtain, but perhaps overall the best in practice.

Hyperinterpolation with the Gauss-Legendre quadrature weights and points is consistently effective, easy to implement, and comparable in quality with the best interpolation schemes. On the other hand, it does need approximately twice as many values of the function being approximated.

Finally among the non-polynomial approximations it is interesting to note the consistently good performance of the scheme of Ganesh, Graham and Sivaloganathan (GGS), in which the polynomials of degree $\leq n$ are augmented by a approximately equal number of non-polynomial basis functions.

**Acknowledgements:** The support of the Australian Research Council is gratefully acknowledged, as is the programming support of Dave Dowsett.

# References

[1] D. L. Berman, *On a class of linear operators*, Dokl. Akad. Nauk SSSR, 85 (1952), pp. 13–16. (Russian), Math. Reviews **14**, 57.

[2] I. K. Daugavet, *Some applications of the Marcinkiewicz-Berman identity*, Vestnik Leningrad Univ., Math., 1 (1974), pp. 321–327.

[3] P. J. Davis, *Interpolation and Approximation*, Blaisdell, New York, 1963.

[4] J. Fliege and U. Maier, *A two-stage approach for computing cubature formulae for the sphere*, Ergebnisberichte Angewandte Mathematik 139T, Universität Dortmund, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany, September 1996.

[5] J. Fliege and U. Maier, *The distribution of points on the sphere and corresponding cubature formulae*, IMA J. Num. Anal., 19 (1999), pp. 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), pp. 1378–1414.

[8] ——, *A new spectral boundary integral collocation method for three-dimensional potential problems*, SIAM J. Numer. Anal., 35 (1998), pp. 778–805.

[9] T. H. Gronwall, *On the degree of convergence of Laplace's series*, Trans. Amer. Math. Soc., 15 (1914), pp. 1–30.

[10] K. Jetter, J. Stöckler, and J. D. Ward, *Error estimates for scattered data interpolation on spheres*, Math. Comp., 68 (1999), pp. 733–747.

[11] Q. T. Le Gia and I. H. Sloan, *The uniform norm of hyperinterpolation on the unit sphere in an arbitrary number of dimensions*, Constructive Approximation, 17 (2000, to appear). http://link.springer.de/link/service/journals/00365/contents/00/10025/.

[12] A. S. Lewis and M. L. Overton, *Eigenvalue optimization*, ACTA Numerica, (1996), pp. 149–190.

[13] C. Müller, *Spherical Harmonics*, vol. 17 of Lecture Notes in Mathematics, Springer Verlag, Berlin, New-York, 1966.

[14] M. L. Overton, *Large-scale optimization of eigenvalues*, SIAM J. Optimization, 2 (1992), pp. 88–120.

[15] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997.

[16] M. Reimer, *Interpolation on the sphere and bounds for the Lagrangian square sums*, Resultate Math., 11 (1987), pp. 144–166.

[17] ——, *Constructive Theory of Multivariate Functions*, BI Wissenschaftsverlag, Mannheim, Wien, Zürich, 1990.

[18] ——, *Quadrature rules for the surface integral of the unit sphere based on extremal fundamental systems*, Mathematische Nachrichten, 169 (1994), pp. 235–241.

[19] ——, *The average size of certain Gram-determinants and interpolation on non-compact sets*, in Multivariate Approximation and Splines (Mannheim, 1966), no. 125 in International Series on Numerical Mathematics, Basel, 1998, Birkhauser, pp. 235–244.

[20] ——, *Hyperinterpolation on the sphere at the minimal projection order*, J. Approx. Theory, 104 (2000), pp. 272–286.

[21] M. Reimer and B. Sündermann, *A Remez-type algorithm for the calculation of extremal fundamental systems on the sphere*, Computing, 37 (1986), pp. 43–58.

[22] I. H. Sloan, *Polynomial interpolation and hyperinterpolation over general regions*, J. Approx. Theory, 83 (1995), pp. 238–254.

[23] I. H. Sloan and R. S. Womersley, *The uniform error of hyperinterpolation on the sphere*, in Advances in Multivariate Approximation, W. Haußmann, K. Jetter, and M. Reimer, eds., vol. 107 of Mathematical Research, Berlin, 1999, Wiley–VCH, pp. 289–306.

[24] ——, *Constructive polynomial approximation on the sphere*, Journal of Approximation Theory, 103 (2000), pp. 91–118.

[25] G. Szegö, *Orthogonal Polynomials*, vol. 23 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, Rhode Island, 4th ed., 1975.

[26] D. L. Williamson, J. B. Brake, J. J. Hack, R. Jakob, and P. N. Swarztrauber, *A standard test set for numerical approximations to the shallow water equations in spherical geometry*, Journal of Computational Physics, 102 (1992), pp. 211–224.

[27] R. S. Womersley, *A continuous minimax problem for calculating minimum norm polynomial interpolation points on the sphere*, ANZIAM J., 42 (E) (2000, to appear).