Limiting Values Under Scaling of the Lebesgue Function for Polynomial Interpolation on Spheres

L. Bos and S. De Marchi

Department of Mathematics and Statistics
University of Calgary
Calgary, Alberta T2N 1N4 CANADA

Dipartimento di Matematica e Informatica
Università di Udine
Via delle Scienze 206, 33100 Udine, ITALIA

Scaling of Interpolation on Spheres

L. Bos
Department of Mathematics and Statistics
University of Calgary
Calgary, Alberta T2N 1N4 CANADA

Phone: (403)220-5202
Fax: (403)282-5150
email: lpbos@math.ucalgary.ca

Abstract

We consider interpolation at points on a $d-$dimensional sphere by spherical harmonics and show that, in the limit, as the points coalesce under an angular scaling, the Lebesgue function of the process converges to that of an associated algebraic interpolation problem for the original angles considered as points in $\mathbb{R}^{d-1}$.

Univariate polynomial interpolation is a classical and much studied subject with many beautiful theoretical results as well as being of some practical importance. In contrast, the multivariate case has been much less studied and many basic questions remain open. A special case of particular interest is interpolation on the sphere and it is the purpose of this work to study one basic property of such interpolants: what happens in the limit as the points coalesce to a single point under an angular scaling. Somewhat unexpectedly the answer (see our main Theorem) is that the interpolation problem becomes one on a certain paraboloid!

In order to explain the details we begin with a general setup for polynomial interpolation in several variables. Let $K \subset \mathbb{R}^d$ be compact. The polynomials of degree n, when restricted to K, form a certain vector space which we will denote by $\mathcal{P}_n(K)$. Let $d_n := \dim(\mathcal{P}_n(K))$ and $\{P_1, \cdots, P_{d_n}\}$ a basis for $\mathcal{P}_n(K)$. Suppose that we are given a set $X = \{x_1, \cdots, x_{d_n}\} \subset K$ of d_n distinct points and a function $f \in C(K)$. The polynomial interpolation problem is then to find $P \in \mathcal{P}_n(K)$ so that

$$P(x_j) = f(x_j) \qquad 1 \le j \le d_n.$$

Writing P in the form

$$P = \sum_{k=1}^{d_n} c_k P_k,$$

this amounts to solving the $d_n \times d_n$ linear system of equations

$$\sum_{k=1}^{d_n} c_k P_k(x_j) = f(x_j) \qquad 1 \le j \le d_n$$

which has a unique solution iff the associated Vandermonde matrix

$$V(x_1, \cdots, x_{d_n}) := [P_k(x_j)]_{1 \le j,k \le d_n}$$

is non-singular. If this is indeed the case we may then form the fundamental Lagrange polynomials

$$\ell_k(x) := \frac{\det V(x_1, \cdots, x_{k-1}, x, x_{k+1}, \cdots, x_{d_n})}{\det V(x_1, \cdots, x_{d_n})} \tag{1}$$

having the property that

$$\ell_k(x_j) = \delta_{jk} \qquad 1 \le j, k \le d_n.$$

This allows the interpolant to be written in the form

$$P(x) = \sum_{k=1}^{d_n} f(x_k) \ell_k(x).$$

The function

$$\Lambda_n(x) := \sum_{k=1}^{d_n} |\ell_k(x)|$$

is known as the Lebesgue function of the process. Its maximum value, $\lambda_n := \max_{x\in K} \Lambda_n(x)$ is referred to as the associated Lebesgue constant and gives the norm of the interpolation projection, $\pi_n : C(K) \to \mathcal{P}_n(K),\ f \to P$, where both spaces are equipped with the supremum norm on K. It is fundamental to the study of the convergence of interpolants (see e.g. [C, pp. 147,212] or [SV, Chapters III and IV]).

Now, as is well-known, polynomial interpolation is well behaved under affine scalings. Specifically, for a scaling factor $a > 0$, if we let $\tilde{K} := aK$, $\tilde{X} := aX$ and $\tilde{\ell}_k$ be the corresponding Lagrange polynomials then clearly

$$\tilde{\ell}_k(ax_j) = \delta_{jk} = \ell_k(x_j)$$

so that, in fact,

$$\tilde{\ell}_k(ax) = \ell_k(x) \qquad (\text{in } \mathcal{P}_n(K)) \tag{2}$$

and hence

$$\tilde{\Lambda}_n(ax) = \Lambda_n(x).$$

In particular,

$$\tilde{\lambda}_n := \max_{y\in\tilde{K}} \tilde{\Lambda}_n(y) = \max_{x\in K} \tilde{\Lambda}_n(ax) = \max_{x\in K} \Lambda_n(x) = \lambda_n.$$

It follows that the Lebesgue constants are in fact *invariant* under affine scalings.

However, such affine scalings are not always what is required in practice. Consider, for example, interpolation of circular data by trigonometric polynomials. In our setting this corresponds to $K = S^1$, the unit circle, $d_n = 2n+1$ and

$$X = \{(\cos(\theta_1), \sin(\theta_1)), \cdots, (\cos(\theta_{2n+1}), \sin(\theta_{2n+1}))\}.$$

The scaling appropriate to such a situation is *angular*, i.e.,

$$\tilde{X} = \{(\cos(a\theta_1), \sin(a\theta_1)), \cdots, (\cos(a\theta_{2n+1}), \sin(a\theta_{2n+1}))\}.$$

But now there is a major difference. A trigonometric polynomial $t(\theta)$ scaled to $t(a\theta)$ is no longer in general even another trigonometric polynomial and thus a formula such as (2) will not be true. Nevertheless, as we shall see, there is a limiting value to each $\tilde{\ell}_k(a\theta)$ as $a \to 0$ and thus also to $\tilde{\lambda}_n$.

Surprisingly, these limiting values are the corresponding entities in an associated *algebraic* interpolation problem, dependent only on the initial values X. The argument is quite simple, provided a careful choice of basis is made. It goes as follows. Any bivariate polynomial $P(x,y)$ of degree $\le n$ may be written in the form

$$P(x,y) = \sum_{k=0}^{n} a_k(x) y^k \tag{3}$$

where $a_k(x)$ is a polynomial of degree $n-k$. On the unit circle, $x^2+y^2 = 1$ or $y^2 = 1-x^2$, and the substitution of this relation into (3) yields

$$P(x,y) = p_1(x) + y p_2(x)$$

for certain polynomials p_1, p_2 of degrees n and $n-1$ respectively. Hence, the monomials

$$\{x^k \,|\, 0 \leq k \leq n\} \cup \{yx^k \,|\, 0 \leq k \leq n-1\}$$

form a basis for $\mathcal{P}_n(K)$. However, for our purposes, it is more convenient to use the basis

$$\{(x-1)^k \,|\, 0 \leq k \leq n\} \cup \{y(x-1)^k \,|\, 0 \leq k \leq n-1\}. \tag{4}$$

Now, on the circle, consider a typical point, scaled by an angular factor of a, $(x, y) = (\cos(a\theta), \sin(a\theta))$. Its coordinates have power series expansions

$$\begin{aligned} x &= \cos(a\theta) = 1 - \frac{a^2\theta^2}{2!} + \frac{a^4\theta^4}{4!} + \cdots \\ y &= \sin(a\theta) = a\theta - \frac{a^3\theta^3}{3!} + \cdots \end{aligned}$$

so that

$$\begin{aligned} (x-1)^k &= (-\frac{a^2\theta^2}{2!} + \frac{a^4\theta^4}{4!} - \cdots)^k \\ &= a^{2k}(-\frac{\theta^2}{2!} + \frac{a^2\theta^4}{4!} - \frac{a^4\theta^6}{6!} + \cdots)^k \\ &= a^{2k}\left(\frac{(-1)^k}{(2!)^k}\theta^{2k} + a^2 R_k(\theta, a)\right), \end{aligned} \tag{5}$$

for some analytic function $R_k(\theta, a)$ whose precise form is unimportant. Similarly,

$$(x-1)^k y = a^{2k+1}\left(\frac{(-1)^k}{(2!)^k}\theta^{2k+1} + a^2 Q_k(\theta, a)\right), \tag{6}$$

again, for some analytic function $Q_k(\theta, a)$.

Hence, when computing $\ell_j(a\theta)$ by means of the expression (1), we see that the column corresponding to the basis function $(x-1)^k$ has a common factor of a^{2k} and that corresponding to $(x-1)^k y$ one of a^{2k+1}. Since this holds for both numerator and denominator, these powers of a simply cancel out. Thus, letting $a \to 0$, we see from (5) and (6) that the column corresponding to $(x-1)^k$ tends to $(-1)^k\theta^{2k}/(2!)^k$ and that to $(x-1)^k y$ tends to $(-1)^k\theta^{2k+1}/(2!)^k$. The factors $(-1)^k/(2!)^k$ are entirely irrelevant and we therefore may conclude that, as $a \to 0$, $\ell_j(a\theta)$ tends to the corresponding Lagrange polynomial for the algebraic interpolation problem associated to the basis

$$\{\theta^{2k} \,|\, 0 \leq k \leq n\} \cup \{\theta^{2k+1} \,|\, 0 \leq k \leq n-1\} = \{\theta^k \,|\, 0 \leq k \leq 2n\}$$

and the points $\theta_1, \theta_2, \cdots, \theta_{2n+1} \in \mathbb{R}$.

An analogous result holds for higher dimensional spheres. Thus, let $K = S^{d-1} \subset \mathbb{R}^d$, the unit sphere. As is well-known,

$$d_n = \binom{n+d-1}{n} + \binom{n+d-2}{n-1}.$$

Now let $x \in \mathbb{R}^{d-1}$ and $y \in \mathbb{R}$ so that $(x, y) \in \mathbb{R}^d$. Then just as in (3), any $d-$variate polynomial $P(x, y)$ of degree $\leq n$ may be written in the form

$$P(x,y) = \sum_{k=0}^{n} a_k(x) y^k$$

where $a_k(x)$ is a $(d-1)-$variate polynomial of degree at most $n-k$. On S^{d-1}, $y^2 = 1 - |x|^2$ so that

$$P(x,y) = p_1(x) + y p_2(x)$$

for certain polynomials p_1, p_2 of degrees at most n and $n-1$ respectively.

Consequently, the monomials

$$\{x^\alpha \,:\, |\alpha| \leq n\} \cup \{y x^\beta \,:\, |\beta| \leq n-1\}$$

form a basis for $\mathcal{P}_n(S^{d-1})$ but, just as before, it is more convenient to use the basis

$$\{(x - e_1)^\alpha \,:\, |\alpha| \leq n\} \cup \{y(x - e_1)^\beta \,:\, |\beta| \leq n-1\}$$

where $e_1 := (1, 0, \cdots, 0) \in \mathbb{R}^{d-1}$.

Suppose then that we are given d_n distinct points $(x^k, y^k) \in S^{d-1}$, $1 \leq k \leq d_n$. We will write each such point in slightly non-standard spherical coordinates:

$$\begin{aligned}
x_2^k &= \cos(\theta_1^k)\sin(\theta_2^k)\\
x_3^k &= \cos(\theta_1^k)\cos(\theta_2^k)\sin(\theta_3^k)\\
&\cdot\\
&\cdot\\
x_{d-1}^k &= \cos(\theta_1^k)\cos(\theta_2^k)\cdots\cos(\theta_{d-2}^k)\sin(\theta_{d-1}^k)\\
x_1^k &= \cos(\theta_1^k)\cos(\theta_2^k)\cdots\cos(\theta_{d-2}^k)\cos(\theta_{d-1}^k)\\
y^k &= \sin(\theta_1^k)
\end{aligned} \tag{7}$$

and then scale the angles θ_j by the factor a. Thus, the points coalesce at $(1, 0, \cdots, 0)$ as $a \to 0$. (Note: this choice is made so that each coordinate has *one* sin factor, which ensures that they converge at the same rate. Other choices of coordinates may not scale all of them in a uniform manner, which will have a pronounced affect on the limit. See the Example below.) We calculate, much as before (suppressing the superscript),

$$\begin{aligned}
x_2 &= \cos(a\theta_1)\sin(a\theta_2)\\
&= (1-\frac{a^2\theta_1^2}{2!}+\frac{a^4\theta_1^4}{4!}+\cdots)(a\theta_2-\frac{a^3\theta^3}{3!}+\cdots)\\
&= a\theta_2+a^3R_2(a,\theta_1,\theta_2)\\
&= a\{\theta_2+a^2R_2(a,\theta_1,\theta_2)\}\\
x_3 &= \cos(a\theta_1)\cos(a\theta_2)\sin(a\theta_3)\\
&= (1-\frac{a^2\theta_1^2}{2!}+\cdots)(1-\frac{a^2\theta_2^2}{2!}+\cdots)(a\theta_3-\frac{a^3\theta_3^3}{3!}+\cdots)\\
&= a\theta_3+a^3R_3(a,\theta_1,\theta_2,\theta_3)\\
&= a\{\theta_3+a^2R_3(a,\theta_1,\theta_2,\theta_3)\}\\
&\text{etc..}
\end{aligned}$$

In general, we have,

$$x_k = a\theta_k + a^3R_k(a,\theta_1,\cdots,\theta_k),\quad 2\le k\le d-1.$$

The first coordinate is somewhat different:

$$\begin{aligned}
x_1-1 &= \cos(a\theta_1)\cos(a\theta_2)\cdot\cdot\cos(a\theta_{d-1})-1\\
&= (1-\frac{a^2\theta_1^2}{2!}+\cdots)\cdots(1-\frac{a^2\theta_{d-1}^2}{2!}+\cdots)-1\\
&= a^2\{\frac{-1}{2!}(\theta_1^2+\cdots+\theta_{d-1}^2)+a^2R_1(a,\theta_1,\cdots,\theta_{d-1})\}.
\end{aligned}$$

Finally,

$$\begin{aligned}
y &= \sin(a\theta_1)\\
&= a\{\theta_1+a^2R_d(a,\theta_1)\}.
\end{aligned}$$

The precise forms of the analytic functions $R_1,\cdots,R_d$ are unimportant.

Again, when substituting these expressions into (1), we find that the common powers of a cancel, and thus in the limit, as $a\to 0$, we replace

$$(x-e_1)^\alpha \quad\text{by}\quad (\theta_1^2+\cdots+\theta_{d-1}^2)^{\alpha_1}\theta_2^{\alpha_2}\cdots\theta_{d-1}^{\alpha_{d-1}}$$

and

$$y(x-e_1)^\beta \quad\text{by}\quad (\theta_1^2+\cdots+\theta_{d-1}^2)^{\beta_1}\theta_1\theta_2^{\beta_2}\cdots\theta_{d-1}^{\beta_{d-1}}.$$

Clearly then, in the limit, we obtain the problem of interpolation at the d_n points

$$\{\theta^k\,|\,1\le k\le d_n\}\subset\mathbb{R}^{d-1}$$

by polynomials of the form

$$p_1(\theta_2,\cdots,\theta_d)+\theta_1p_2(\theta_2,\cdots,\theta_d) \tag{8}$$

where $\deg(p_1) \le n$, $\deg(p_2) \le n-1$ and $\theta_d = \theta_1^2 + \theta_2^2 + \cdots + \theta_{d-1}^2$. Let A_n denote the space of such polynomials, i.e.,

$$A_n := \{p_1(\theta_2, \cdots, \theta_d) + \theta_1 p_2(\theta_2, \cdots, \theta_d) \,|\, \deg(p_1) \le n,\ \deg(p_2) \le n-1,\ \theta_d = \theta_1^2 + \cdots + \theta_{d-1}^2\}.$$

Since, p_1, p_2 are polynomials in $d-1$ variables, it follows that

$$\begin{aligned}
\dim(A_n) &\le \dim(\mathcal{P}_n(\mathbb{R}^{d-1})) + \dim(\mathcal{P}_{n-1}(\mathbb{R}^{d-1})) \\
&= \binom{n+d-1}{n} + \binom{n-1+d-1}{n-1} \\
&= d_n.
\end{aligned}$$

We claim that, in fact, $\dim(A_n) = d_n$ and hence this interpolation problem is well-posed. To show this, it suffices to show that

$$p_1(\theta_2, \cdots, \theta_{d-1}, \theta_1^2 + \cdots + \theta_{d-1}^2) + \theta_1 p_2(\theta_2, \cdots, \theta_{d-1}, \theta_1^2 + \cdots + \theta_{d-1}^2) \equiv 0 \qquad (9)$$

for all $\theta_1, \cdots, \theta_{d-1} \in \mathbb{R}$ iff $p_1 \equiv 0$ and $p_2 \equiv 0$. So suppose that (9) holds. Since this is an algebraic identity it also holds for all $\theta_1, \cdots, \theta_{d-1} \in \mathbb{C}$. In particular, it holds for $\theta_1 = it(\theta_2^2 + \cdots + \theta_{d-1}^2)^{1/2}$, $t \in \mathbb{R}$, for which $\theta_1^2 + \cdots + \theta_{d-1}^2 = (1-t^2)(\theta_2^2 + \cdots + \theta_{d-1}^2)$. Hence,

$$\begin{aligned}
&p_1(\theta_2, \cdots, \theta_{d-1}, (1-t^2)(\theta_2^2 + \cdots + \theta_{d-1}^2)) \\
&\qquad + it(\theta_2^2 + \cdots + \theta_{d-1}^2)^{1/2} p_2(\theta_2, \cdots, \theta_{d-1}, (1-t^2)(\theta_2^2 + \cdots + \theta_{d-1}^2)) \equiv 0.
\end{aligned}$$

Equating real and imaginary parts, we have

$$p_1(\theta_2, \cdots, \theta_{d-1}, (1-t^2)(\theta_2^2 + \cdots + \theta_{d-1}^2)) \equiv 0$$

and

$$p_2(\theta_2, \cdots, \theta_{d-1}, (1-t^2)(\theta_2^2 + \cdots + \theta_{d-1}^2)) \equiv 0.$$

But since $t \in \mathbb{R}$ is arbitrary it follows that $p_1 \equiv 0$ and $p_2 \equiv 0$, as claimed.

There is an interesting geometric interpretation of the space A_n. By definition it consists of polynomials of the special form (8), restricted to the *paraboloid* $\theta_d = \theta_1^2 + \cdots + \theta_{d-1}^2$. But the dimension of the space of *all* polynomials in $\theta_1, \cdots, \theta_d$, restricted to this paraboloid is also d_n (for the same reason as for the sphere). Hence, letting

$$Q^{d-1} := \{(\theta_1, \cdots, \theta_d) \,|\, \theta_d = \theta_1^2 + \cdots + \theta_{d-1}^2\} \subset \mathbb{R}^d$$

be this paraboloid, we have

$$A_n = \mathcal{P}_n(Q^{d-1}).$$

In summary, we have established the following theorem giving the *"limiting value under scaling"* for polynomial interpolation on spheres.

Theorem. Suppose we are given d_n distinct points on the sphere, $(x^k, y^k) \in S^{d-1}$, $1 \le k \le d_n$. Consider two interpolation problems associated to these points:

(1) *Spherical Interpolation.* Interpolate a function $f \in C(S^{d-1})$ at the points (x^k, y^k) by a polynomial $p \in \mathcal{P}_n(S^{d-1})$. Denote the fundamental Lagrange polynomials by $\ell_k(x)$.

(2) *Paraboloidal Interpolation.* First write the points in spherical coordinates as in (7). Interpolate a function $g \in C(Q^{d-1})$ at the d_n points $(\theta^k, \theta_d^k) \in Q^{d-1}$ with $\theta_d = \theta_1^2 + \cdots + \theta_{d-1}^2$ by a polynomial $p \in \mathcal{P}_n(Q^{d-1})$. Denote the fundamental Lagrange polynomials by $\tilde{\ell}_k(\theta)$ (if they exist, i.e. the problem is *unisolvent*).

Let us denote by $x(a)$ the point obtained by scaling the angles in (7) by the factor a and by ℓ_k^a the fundamental Lagrange polynomial associated to all the given points on the sphere scaled by a in this manner. Then

$$\lim_{a \to 0} \ell_k^a(x(a)) = \tilde{\ell}_k(\theta).$$

The geometry underlying the Theorem is illustrated in Figure 1, below, for the case $d = 2$.

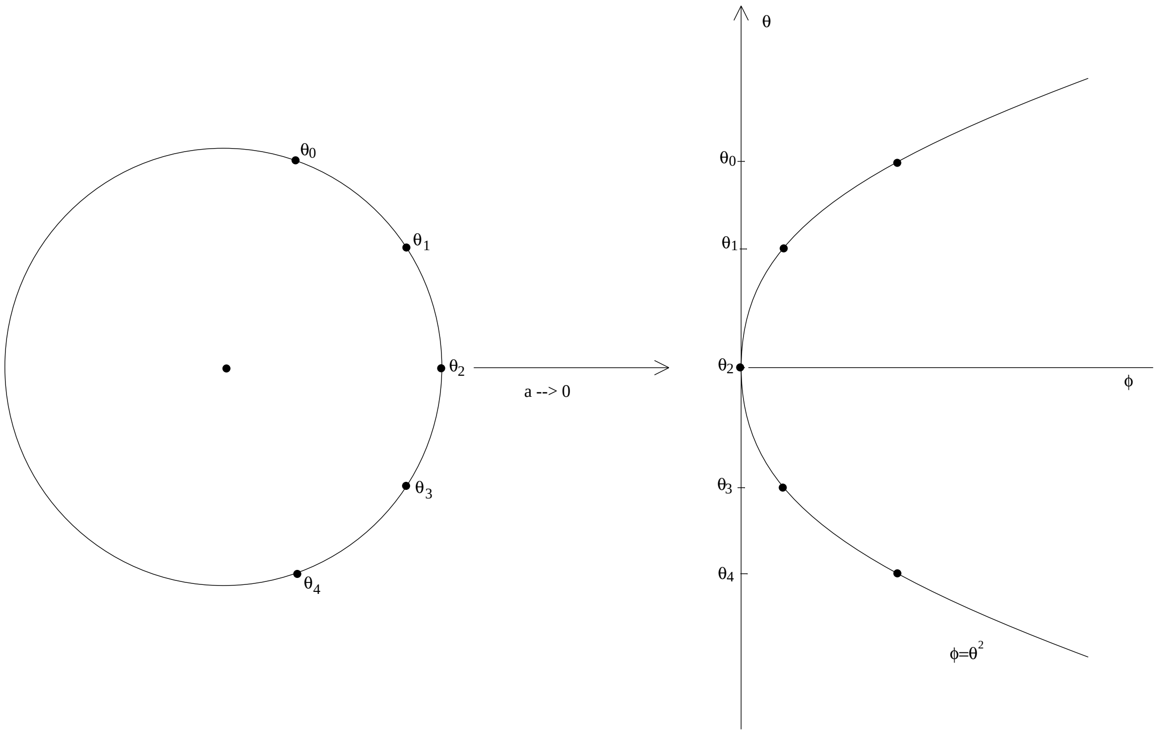

Figure 1. The Limiting Process

Corollary. Suppose that K is a compact subset of S^{d-1} and that the set of points $X \subset K$. Just as before let $\lambda_n := \max_{x \in K} \Lambda_n(x)$ denote the associated Lebesgue constant. Further, let $\tilde{K}$ and $\tilde{X}$ denote the angular scalings of these sets by means of the spherical coordinates (7), and $\tilde{\lambda}_n$ the corresponding Lebesgue constant. Then, if we let K_θ be the image of K under the inverse of the mapping $\theta \to x$ given by (7), the limit of $\tilde{\lambda}_n$ as $a \to 0$ is the Lebesgue constant associated to the problem (2) in the Theorem on the set K_θ.

In general, the limiting values of the Lagrange polynomials depends strongly on how the points coalesce. We give a simple example illustrating this fact.

Example. Consider linear interpolation at four points on the sphere S^2. The spherical coordinates (7) specialize to

$$\begin{aligned} z &= \sin(\phi) \\ x &= \cos(\phi)\cos(\theta) \\ y &= \cos(\phi)\sin(\theta) \end{aligned}$$

and the limit of the scaled interpolation problem corresponds to interpolation at the points (θ, ϕ) by a polynomial in the linear span of $\{1, \theta, \phi, \theta^2 + \phi^2\}$.

If, on the other hand, we were to use the more standard spherical coordinates

$$\begin{aligned} z &= \cos(\phi) \\ x &= \sin(\phi)\cos(\theta) \\ y &= \sin(\phi)\sin(\theta) \end{aligned}$$

the limit would correspond to interpolation from the linear span of $\{1, \phi, \theta\phi, \phi^2\}$, as is easily seen from the Taylor expansions of the coordinate functions.

We conclude with several numerical examples illustrating the convergence for points on the circle. Figure 2 is of 5 equally spaced points on $-\pi/4 \leq \theta \leq \pi/4$ and Figure 3 is of 7 points chosen at random. Although, the convergence is monotone in both cases, Figure 4 shows that this is not *always* true.

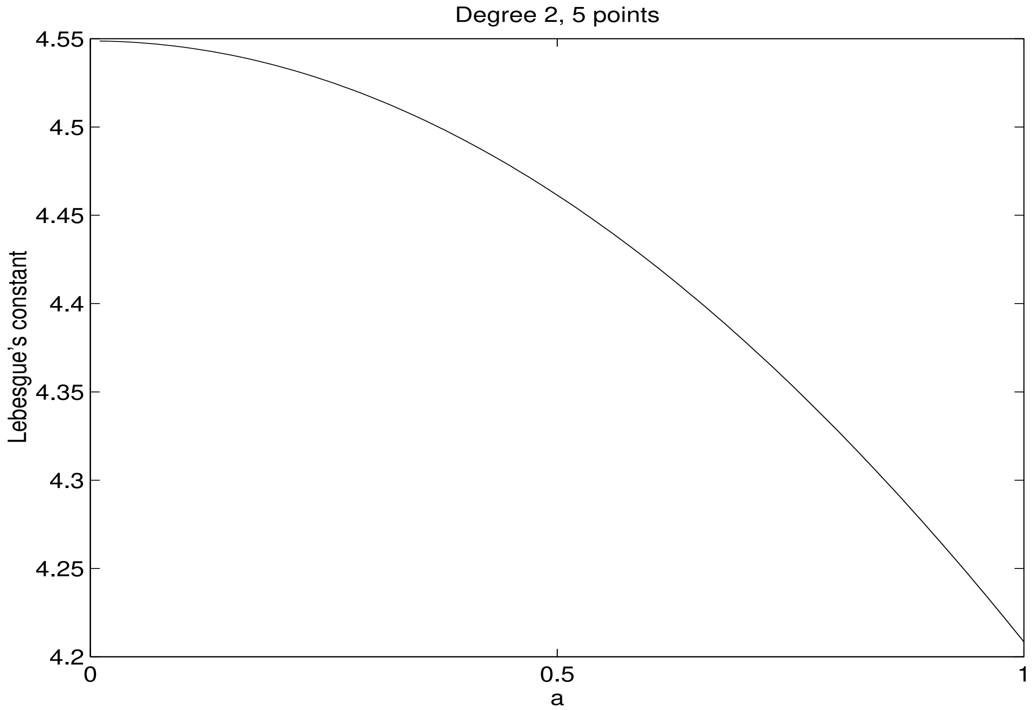

Figure 2. The Behaviour of the Lebesgue Function
for 5 equally spaced points

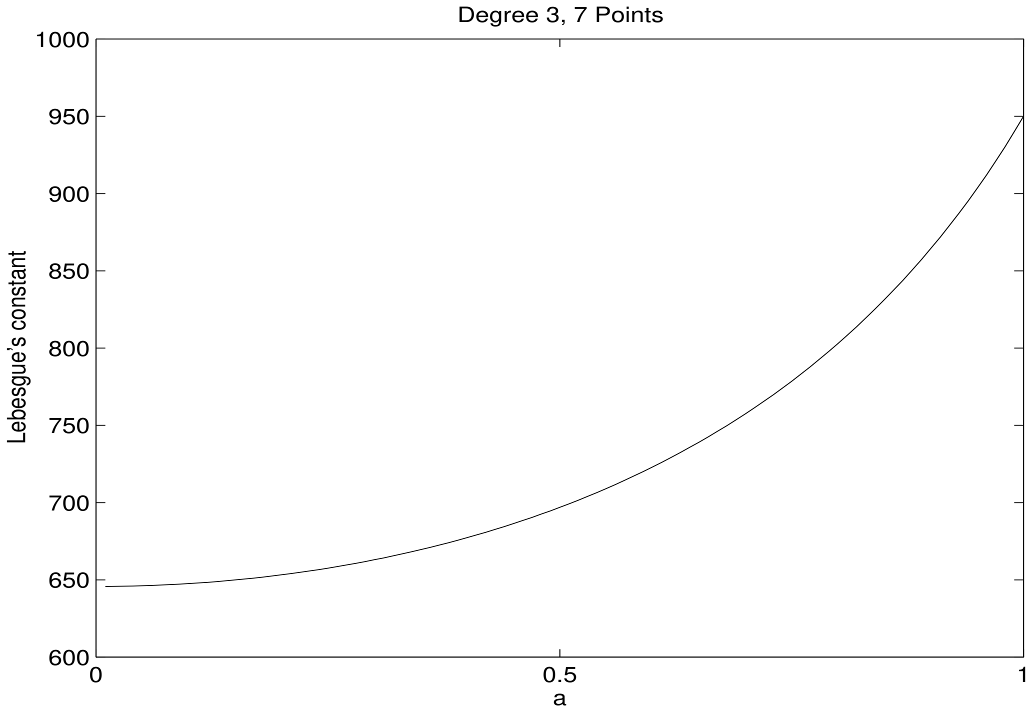

Figure 3. The Behaviour of the Lebesgue Function
for 7 points chosen at random

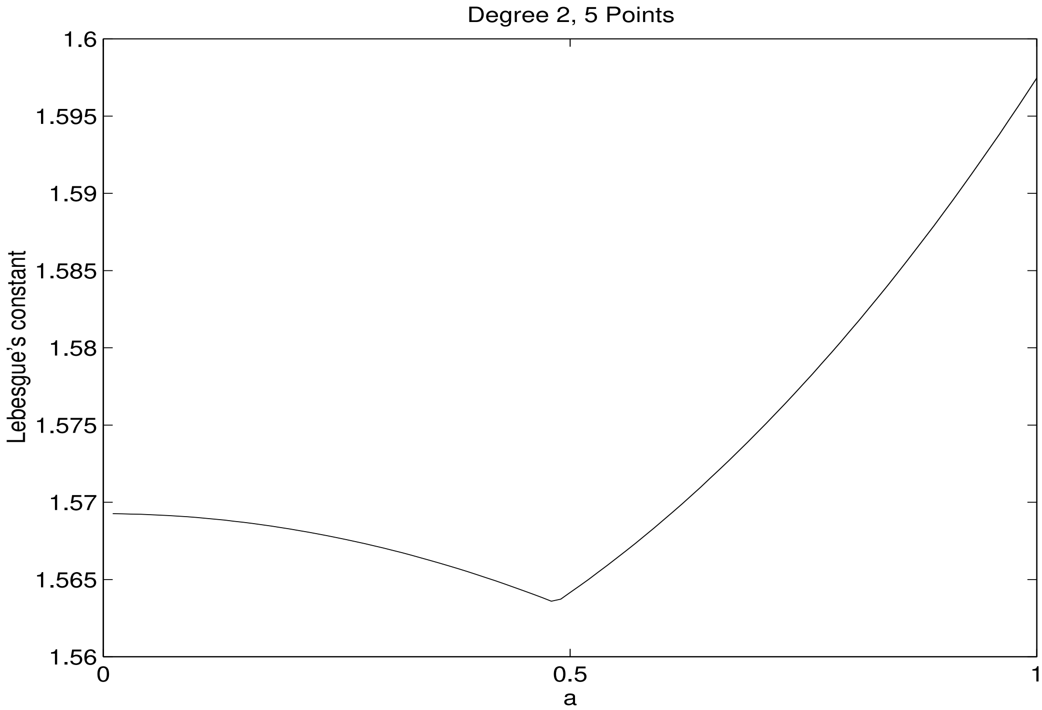

Figure 4. The Behaviour of the Lebesgue Function
for 5 Chebychev points

REFERENCES

[C] E. W. Cheney, "Introduction to Approximation Theory", McGraw-Hill, 1966.

[SV] J. Szabados and P. Vértesi, "Interpolation of Functions", World Scientific, Teaneck NJ, 1990.