Abstract

Single-domain spectral methods have been largely restricted to tensor product bases on a tensor product grid. To break the “tensor barrier”, we studied approximation in two idealized families of domains. One family is bounded by a “squircle”, the zero isoline of $B(x, y) = x^{2\nu} + y^{2\nu} − 1$. The boundary varies smoothly from a circle [$\nu = 1$] to the square [$\nu = \infty$]. The other family is bounded by a “perturbed quadrifolium”, the plane algebraic curve $\delta(x^2 + y^2) − ((x^2 + y^2)^3 - (x^2 - y^2)^2)$; this varies smoothly from the singular, self-intersecting curve known as the quadrifolium to a circle as $\delta$ varies from zero to infinity. We compared two different bivariate polynomial bases, truncating by total degree. Zernike polynomials are natural for the disk; tensor products of Chebyshev polynomials are equally sensible for the square. Both yield an exponential rate of convergence for our non-tensor, neither disk-nor-square domains; indeed, the Chebyshev basis worked well for the disk and the Zernike polynomials were good for the square. The expected differences due to numerical ill-conditioning did not emerge, much to our surprise. The price for the nontensor domain was that hyperinterpolation was necessary, that is, least squares fitting with more interpolation constraints than unknowns. Denoting the number of interpolation points by $P$ and the basis size by $N$, a ratio of $P/N$ around two to three was optimum while $P$ near one was very inaccurate. A uniform grid, truncated to include only those points within the squircle or other boundary curve, was satisfactory even without interpolation points on the boundary (although boundary points are a cost-effective improvement). Interpolation costs were greatly reduced by exploiting the invariance of the squircle-bounded and perturbed-quadrifolium domains to the eight element $D_4$ dihedral group.

Keywords: pseudospectral; hyperinterpolation; Chebyshev polynomials;
1. Introduction

Spectral methods have been hitherto limited to tensor product domains such as rectangles, disks and triangles. A major research frontier is to extend single domain spectral methods to more complicated geometries.
Table 1: Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B(x, y)$</td>
<td>Boundary function: its zero isoline is the boundary</td>
</tr>
<tr>
<td>$B_j(x)$</td>
<td>Generic basis functions</td>
</tr>
<tr>
<td>$D$</td>
<td>Total degree of a polynomial; total degree($x^m + y^n$) = $m = n$</td>
</tr>
<tr>
<td>$M$</td>
<td>Numbers of points in $x$ (or $y$) for a tensor product grid on the square</td>
</tr>
<tr>
<td>$N$</td>
<td>Total number of basis functions</td>
</tr>
<tr>
<td>$P$</td>
<td>Total number of sampling (hyperinterpolation) points</td>
</tr>
<tr>
<td>$q$</td>
<td>Parameterization for the squircle: $r = q(\theta; \nu) \equiv {\cos^{2\nu}(\theta) + \sin^{2\nu}(\theta)}^{-1/(2\nu)}$</td>
</tr>
<tr>
<td>$r$</td>
<td>Radial coordinate of a polar coordinate system</td>
</tr>
<tr>
<td>$R$</td>
<td>Radial coordinate in the transformed polar coordinates</td>
</tr>
<tr>
<td>$S(\nu)$</td>
<td>Radial scaling constant for the Zernike polynomials</td>
</tr>
<tr>
<td>$T_n(x)$</td>
<td>Chebyshev polynomial of degree $n$</td>
</tr>
<tr>
<td>$Z_m^n(x, y)$</td>
<td>Zernike polynomial of angular wavenumber $m$ and total degree $n$</td>
</tr>
<tr>
<td>$Z_m^n(r, \theta)$</td>
<td>Zernike disk polynomial expressed in polar coordinates</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Relative shape parameter, $\epsilon$ divided by average grid spacing</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Absolute shape parameter (inverse width)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$P/N$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Exponent in the definition of the squircle</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Location of one endpoint in Chebyshev Extension</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Angular coordinate in polar coordinates</td>
</tr>
<tr>
<td>$\aleph$</td>
<td>Scaling constant in analysis of Zernike scaling</td>
</tr>
</tbody>
</table>

In this, the second part of a series, we wrestle with a number of issues that must be sorted out before proceeding to the ambitious PDE computations of Part 3 and beyond.

Because passing the “tensor barrier” is very difficult, we study two families of simply connected domains whose boundaries are a squircle and a perturbed quadrifolium, respectively.

Tensor products of Chebyshev polynomials in Cartesian coordinates, $T_n(x)T_n(y)$, are canonical basis functions for the square; Zernike polynomials $Z_m^n(x, y)$ are canonical for the disk. A major question addressed here is: What is the best polynomial basis for a geometry that is neither square nor disk?

Although our strategy is “meshless”, which is the accepted shorthand for “subdomain-less, global expansion” methods, the choice of grid is non-trivial when the geometry is non-tensor. Here, we will employ the truncated uniform square grid both with and without boundary points. Grids with Chebyshev-like boundary-biasing are discussed in Part 3; other future work will examine radially-stretched Chebyshev grids for star-like domains, conformal mapping and so on. None of these other options has the simplicity of a uniform grid, and it is worthwhile to see how far the simplest option can be pushed.

Besides interpolation, we shall also employ “hyperinterpolation”. This was coined by Ian Sloan to describe approximations where the coefficients of the basis functions are computed by numerical quadratures using a set of points whose cardinality exceeds that of the basis set [33]. Here, we do not use quadratures,
but rather least-squares approximation with more interpolation conditions than basis coefficients.

**Definition 1 (Least-squares hyperinterpolation).** Suppose an approximation \( f_N \) has the form

\[
f_N(x) \equiv \sum_{k=1}^{\infty} a_k B_k(x)
\]

and determines the coefficients \( a_j \) by solving the matrix equation

\[
\vec{G}\vec{a} = \vec{f}
\]

where the elements of the \( N \)-dimensional vector \( \vec{a} \) are the spectral coefficients \( a_n \), the elements of the \( P \)-dimensional vector \( \vec{f} \) are the “samples” or “grid point values” of the function \( f(x) \) being approximated at a set of sampling points \( x_j \), \( f(x_j) \), and the elements of the “interpolation” or “Vandermonde” matrix are

\[
G_{jk} = B_k(x_j)
\]

If the number of points \( P \) equals \( N \), the number of basis functions, the result is “interpolation”. If \( P < N \) so that the matrix system is underdetermined, the approximation is “hypointerpolation”. If \( P > N \), the result is “hyperinterpolation”.

There are useful strategies that we will briefly discuss, but not apply here either because of difficulties with this geometry or to avoid prolonging this article to excessive length. Omitted strategies include:

1. A truncated uniform hexagonal grid [22, 3]
2. Fourier Extension [1, 24, 8, 10]
4. Mapping the domain to a square or disk through an analytic change of coordinates [31, 29]
5. Polynomials constructed to be orthogonal on the squircle [20, 21].

Thus, our analysis of the squircle and perturbed-quadrifolium will be incomplete. We shall nevertheless present a variety of new ideas that do break the “rectangularity barrier”.

2. Preliminaries: Boundary Curves and Basis Truncation

2.1. Two Simply-Connected Domains Bounded by Plane Algebraic Curves

A “squircle” is a mathematical shape with properties between those of a square and those of a circle. It is defined in implicit form by the following equation,

\[
x^{2\nu} + y^{2\nu} = 1
\]
where \( \nu \) is restricted to an integer here, though this is not strictly necessary. As \( \nu \) increases, the curve becomes more and more like a square with slightly rounded corners; the limit as \( \nu \to \infty \) is a square. In polar coordinates, the parametric equation is

\[
r = q(\theta; \nu) \equiv \left( \cos^{2\nu}(\theta) + \sin^{2\nu}(\theta) \right)^{-1/(2\nu)}
\]  

Although we shall restrict attention to equal axes, polynomial hyperinterpolation and RBF interpolation apply equally well (and badly) to superellipses, and not merely to squircles. The left panel of Fig. 1 shows three members of the squircle family.

\[
B(x, y) = \delta(x^2 + y^2) - ((x^2 + y^2)^3 - (x^2 - y^2)^2)
\]  

where the quadrifolium is the limit \( \delta \to 0 \) and a circle of radius \( \delta^{1/4} \) is the limit \( \delta \to \infty \). The curve can also be parameterized as \( r(\theta) \) where

\[
r = (1/2) \sqrt{1 + \cos(4\theta) + \sqrt{16\delta + 1 + 2\cos(4\theta) + \cos^2(4\theta)}}
\]  

The right panel of Fig. 1 illustrates a typical member of the perturbed quadrifolium family.
2.2. Basis Truncation

For reasons explained in Part I [25], we here adopt a total degree truncation for polynomial bases. For a tensor product Chebyshev basis, a term like $T_m(x)T_n(x)$ is kept in the truncation of total degree $D$ only if

$$m + n \leq D$$

(8)

Although this has been championed in association with the grid known as the “Padua points”, total degree truncation is otherwise rare for Chebyshev polynomials. However, total degree truncation is natural and normal when Zernike polynomials are applied on the disk as explained in [14]. Also, to compare different basis sets, it is helpful to apply the same truncation to each.

3. Truncated Uniform Grids With and Without Points on the Boundary

As explained in the next section, optimal grids, especially for polynomials, are boundary-concentrated like a Chebyshev grid in one dimension. Unfortunately, creating a non-uniform grid in complicated geometry is not easy. We therefore describe, in this section and the next, the two types of grids we used. In an appendix, we explain why a hexagonal grid, so useful in other applications, is unhelpful here.

3.1. Truncated uniform square grid without boundary points

The uniform square grid is constructed in two stages. First, construct a uniform grid for the unit square by choosing an integer $M$ and then defining

$$x_j^s = -1 + 2(j - 1)/(M - 1); \quad y_k^s = -1 + 2(k - 1)/(M - 1)$$

(9)

Second, test each point on the uniform square grid for membership in the interior of the domain by evaluating the function that implicitly defines the squircle,

$$B(x, y) \equiv 1 - x^{2\nu} - y^{2\nu}$$

(10)

A point is on the interior if and only if $B(x, y) > 0$. A counter is increased by one every time a point passes the test so that at the end of the double loop the counter is equal to $N_I$, the number of interior points.

Gridgeman shows that the area of a domain bounded by a squircle is [23]

$$A = 4 \int_0^1 (1 - x^{2\nu})^{1/(2\nu)} dx = \frac{\left(\frac{\Gamma(\frac{1}{2})}{\nu \Gamma(1/\nu)}\right)^2}{\nu \Gamma(1/\nu)}$$

(11)

In the limit that the number of interpolation points is large, the number of included points will asymptote to $A(\nu)/4$. (The area of the embedding square is 4.)
3.2. Boundary points

When there are no points on the boundary itself, all interpolation methods become extrapolation methods in the sense that the approximation must be evaluated outside the convex hull of the interpolation points. Obviously, as one moves farther and farther outside the region of the samples of \( f(x, y) \), the error in approximating \( f(x, y) \) will grow rapidly.

Therefore is highly desirable to place interpolation points on the boundary itself. This entails some difficulties:

1. One needs either an explicit parametric representation of the boundary or a spectrally-accurate method for calculating the zero contour line of the implicit boundary function \( B(x, y) \).
2. Boundary-gridding is a second, separate stage; it is additional work superimposed on the task of defining grid points for the interior.
3. The choice of \( N_B \), the number of boundary points, is another numerical free parameter.

For the squircle, both implicit and explicit parametric representations are available, so the first difficulty disappears. For the square, extending a tensor product \( M \times M \) grid to the boundaries automatically yields \( 4M - 4 \) boundary points, or in other words, using \( N_I = (M - 2)^2 \) as the total number of interior grid points,

\[
N_B \sim 4 \sqrt{N_I}, \quad N_I \gg 1 \tag{12}
\]

This provides a useful guideline for the squircle and more complicated geometries, however, extensive experimentation may be necessary to optimize the number of boundary points.

4. Blow-Up and Scaling in Zernike polynomials

4.1. Review: Properties of Zernike polynomials

Frits Zernike (1888–1966) won the Nobel prize for physics in 1953 for his invention of the phase contrast microscope. In related work, he observed that for optics it was convenient to expand functions on the disk as a Fourier series in angle combined with one-sided Jacobi polynomials in radius [14].

A squircle-bounded domain is not a disk. However, there is no canonical basis for a squircle-disk. Since the circle is one limit of the squircle, a Zernike basis is an obvious choice for experiments. The square is also a limit of the squircle, and thus we experiment with tensor product Chebyshev bases, too.

Zernike polynomials fall into two classes which are respectively antisymmetric and symmetric with respect to \( \theta = 0 \) where \((r, \theta)\) are the usual polar coordinates:

\[
\begin{align*}
Z_n^{-m}(r, \theta) &= R_n^m(r) \sin(m \theta) \\
Z_n^m(r, \theta) &= R_n^m(r) \cos(m \theta)
\end{align*}
\]

\[
\begin{cases}
  n \geq m \geq 0, \quad n = m, |m+2, m+4, \ldots \tag{13}
\end{cases}
\]
The radial functions are polynomials of degree \( n \) defined explicitly as

\[
R_m^n(r) = \begin{cases} 
\sum_{k=0}^{(n-m)/2} \frac{(-1)^k(n-k)!}{k!(1/2)(n+m-k)!((1/2)(n-m)-k)!} r^{n-2k} & n - m \text{ even} \\
0 & n - m \text{ odd}
\end{cases}
\]  

(14)

The radial Zernike polynomials may also be defined in terms of Jacobi polynomials by

\[
R_m^n(r) = (-1)^{(n-m)/2} r^m P_{(n-m)/2}^{m,0}(1 - 2r^2)
\]

(15)

where \( P_{n}^{\alpha,\beta}(r) \) is the usual Jacobi polynomial [30].

**Theorem 1.**

1. \( Z_m^n \) is defined only when \( |m| \leq n \).
2. \( Z_m^n \) is a polynomial of degree \( n \) in radius.
3. \( Z_m^n \) is nonzero only when \( n - m \) is even.
4. There are \((2n + 1)\) possible \( m \) values for a given \( n \), \( m = -n, \ldots, 0, \ldots, n \), but only \((n + 1)\) nonzero polynomials of subscript \( n \).

[14, 19]

This basis set has been independently rediscovered several times. The label “Zernike” is universal in optics, but the name “one-sided Jacobi” is customary in fluid mechanics. The reason for the adjective “one-sided” is that the Jacobi polynomials \( P_{\alpha,\beta}^n(z) \) are orthogonal on the interval \( z \in [-1, 1] \) with respect to the weight function \((1-z)^{\alpha} \). Thus, substituting \( z = 1 - 2r^2 \), the set \( P_{(n-m)/2}^{m,0}(1-2r^2) \) for different \( n \) are polynomials of degrees \((n-m)\) in \( r \), orthogonal on \( r \in [0, 1] \) with a weight function of \( r^{2m+1} \). The Legendre polynomials \( P_n(x) \) are the special case \( m = 0 \), that is, \( R_n^0(r) = (-1)^n P_n(1-2r^2), n = 0, 1, 2, \ldots \).

An important property is the following.

**Theorem 2 (Zernike Polynomials are Cartesian Polynomials).** The Zernike polynomial \( Z_m^n(r, \theta) \) is a bivariate polynomial of total degree \( n \) in the Cartesian coordinates \((x, y)\). [14, 19]

Thus, in infinite precision arithmetic and with total degree truncation, a set of Zernike polynomials spans the same subspace as products of Chebyshev polynomials in \( x \) and \( y \). The subtle question is: How large are the Zernike/Chebyshev differences for approximation in finite precision arithmetic on irregular domains?

### 4.2. Scaling of Zernike polynomials

The points \((\pm 1, 0)\) and \((0, \pm 1)\) are on the squircle for all values of the exponent \( \nu \geq 1 \). Consequently, the smallest square which completely encloses the squircle is the unit square. Since the natural domain of a tensor product Chebyshev basis is the unit square, our Chebyshev basis will be the products of the usual Chebyshev polynomials without any scaling for different \( \nu \).

The situation is more complicated with Zernike polynomials because their natural home is the unit disk. The Zernike polynomials grow exponentially
with distance from the origin \( r = \sqrt{x^2 + y^2} \) when \( r > 1 \) as illustrated in Fig. 2. However, along the diagonal lines \( y = \pm x \), the range of the radial coordinate exceeds one, reaching a maximum of \( \sqrt{2} \) for the limiting case of a square; the radial interval for the square is shown in the graph. It follows that use of unscaled Zernike polynomials is disastrous. Therefore, we employed the Zernike polynomials only as the scaled basis

\[
B_{mn}(x, y) = Z_m^n(S(\nu)x, S(\nu)y)
\]

where the scaling factor is chosen so that \( (S(\nu)x, S(\nu)y) \) has unit distance from the origin in the corners,

\[
S = 2^{(1-\nu)/(2\nu)}
\]

The effects of \( S \) are best illustrated by an example. Let \( f(x, y) \equiv 1/(2 + x^2 + y^2) \). In polar coordinates \( (r, \theta) \), this function is independent of angle, simplifying to \( f(r, \theta) = 1/(2 + r^2) \). Because of the scaling, the approximation problems must be expressed in terms of the scaled radial coordinate \( \rho \equiv S(\nu)r \) where \( \rho \in [0, 1] \).

Then

\[
f = \frac{1}{2 + \rho^2/S(\nu)^2} = \frac{S(\nu)^2}{2S(\nu)^2 + \rho^2} = \sum_{n=0}^{\infty} b_{2n} Z_{2n}^0(\rho, \theta) = \sum_{n=0}^{\infty} b_{2n} R_{2n}^0(\rho) = \sum_{n=0}^{\infty} b_{2n} (-1)^n P_n(1 - 2\rho^2)
\]

where the \( Z_{2n}^0(\rho, \theta) \) are the Zernike polynomials that are independent of angle \( \theta \) and \( R_{2n}^0(\rho) \) are the radial factors of these Zernike polynomials. These furthermore are equal to the usual Legendre polynomials \( P_n(z) \) except for changes of sign and argument as shown.

No simple analytic formula for the Zernike coefficients is known, but theory says that both the Chebyshev and Zernike coefficients will fall proportionally to a factor \( q^n \) which is identical for both types of expansions, and indeed for the coefficients of any series of Jacobi polynomials. This vast family of orthogonal polynomials includes both the Chebyshev and Zernike polynomials as special cases. The Chebyshev coefficients of a reciprocal quadratic function are given by [26, 7]

\[
\frac{N^2}{x^2 + N^2} = \frac{8}{\sqrt{N^2 + 1}} + \frac{2N}{\sqrt{N^2 + 1}} \sum_{n=1}^{\infty} (-1)^n q^n T_{2n}(x), \quad x \in [-1, 1]
\]
where
\[ q \equiv 2N^2 + 1 - 2N\sqrt{N^2 + 1} \] (24)

Here \( N = \sqrt{2}S(\nu) \) and thus
\[ q \equiv 4S(\nu)^2 + 1 - 2\sqrt{2}S(\nu)\sqrt{2S(\nu)^2 + 1} \] (25)

Table 2 lists the numerical values for \( S(\nu) \) and \( q(\nu) \) for various \( \nu \).

Note that the series involve only even degree polynomials so that when the series is truncated at \( n = N \), the degree of the approximation is \( 2N \). The factor \( q(\nu) \) grows as \( \nu \) increases, confirming that Zernike polynomials are indeed at their best on the disk.

Table 2: Scaling factors and \( q(\nu) \) for Zernike expansions

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>( S(\nu) )</th>
<th>( q(\nu) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [disk]</td>
<td>1</td>
<td>0.101</td>
</tr>
<tr>
<td>2</td>
<td>0.841</td>
<td>0.133</td>
</tr>
<tr>
<td>3</td>
<td>0.794</td>
<td>0.145</td>
</tr>
<tr>
<td>4</td>
<td>0.771</td>
<td>0.151</td>
</tr>
<tr>
<td>5</td>
<td>0.758</td>
<td>0.155</td>
</tr>
<tr>
<td>( \infty ) [square]</td>
<td>( 1/\sqrt{2} = 0.707 )</td>
<td>( 3 - 2\sqrt{2} = 0.172 )</td>
</tr>
</tbody>
</table>

Lastly, it is easy to exploit group symmetry using Zernike polynomials with a great reduction in computational cost, even for unsymmetric problems. This is a major theme of our earlier article [25].

5. Spectral Extension: Definition & Overview

Fourier series and also the classical orthogonal polynomial series furnish least-squares solutions to approximation problems only in association with the unique canonical intervals and quadrature weight functions associated with each type of series. Trigonometric series and trigonometric interpolation are optimized for the approximation of periodic functions on domains that are the tensor product of the “one-torus”, which is the interval \( x \in [0, 2\pi] \) with implicit periodicity in \( x \) of period \( 2\pi \). A Chebyshev series is the least squares approximation on \( x \in [-1, 1] \) with the weight function \( w(x) \equiv 1/\sqrt{1-x^2} \). And so on. Through the trivial linear change of coordinates, \( y = (b+a)/2 + (b-a)x/2 \), an arbitrary interval \( y \in [a,b] \) can be mapped into the canonical Chebyshev and Legendre polynomial interval, \( x \in [-1,1] \). When asserting that each type of series is associated with a “unique canonical interval and weight function”, we mean “modulo such simple affine mappings”.

A tensor product of Fourier or Chebyshev polynomial basis functions, applied to a non-tensor domain, is necessarily an “extension” beyond the usual geometry associated with these functions. But such “extensions” can also arise from other motivations.
Fourier series, for example, are associated with a canonical uniform grid as opposed to the highly nonuniform, concentrated-near-the-endpoints Chebyshev grid. This has motivated “Fourier Extension”, which is to approximate a non-periodic function on the “physical” interval \( x \in [-L, L] \) where \( L < \pi \) by means of a Fourier series on the larger “computational” domain \( x \in [-\pi, \pi] \).

One complication with the Chebyshev basis when applied to a square or a domain bounded by a circle is that the physical domain is smaller than the usual domain for the basis. The Chebyshev series will converge equally fast over the entire square because the Chebyshev functions oscillate uniformly over this region. In the limit of a circle-bounded domain, we are, whether we want to or not, approximating \( f(x, y) \) over the square domain from information only from points in a smaller domain. Fourier Extension has been studied in [8, 9, 24, 16, 15, 18, 4, 17]. With a Fourier basis, the interpolation matrix becomes increasingly ill conditioned as the interpolation interval shrinks and accuracy also deteriorates. Even so, excellent results have been achieved by Huybrechs and Adcock [1, 2, 24], Lyon [28, 27], and others cited there and in [8, 9, 10].

This previous work and present needs motivate the following.

**Definition 2 (Spectral Extension).** An approximation using a set of basis functions \( \{B_j\} \) is said to be a “Spectral Extension” when applied to a domain which is different from the usual canonical domain for that basis; this implies that there are regions within the canonical domain where the extended problem has no interpolation points or which is excluded from the integral inner product associated with the basis.

This very general but somewhat abstract definition is a skeleton which we shall now flesh out with examples.

### 6. Extension in One Dimension

The analogous one-dimensional Chebyshev extension is to define a set of interpolation points restricted to a domain \( x \in [-\chi, 1] \) where \( \chi < 1 \) while using standard Chebyshev basis. For example,

\[
f(x) \approx \sum_{n=0}^{N} a_n T_n(x), \quad x \in [-1, 1]
\]

(26)

where the coefficients are determined by interpolation at points such as

\[
x_j^{\text{Chebyshev-like}} = -\chi + \frac{1 + \chi}{2} \left[ 1 + \cos \left( \pi \frac{[2j-1]}{(2N+2)} \right) \right], \quad j = 1, 2, \ldots (N+1)
\]

(27)

or the uniform grid

\[
x_j^{\text{uniform}} = -\chi + (1 + \chi)(j-1)/(N-1), \quad j = 1, 2, \ldots N
\]

(28)
Fig. 3 schematically compares the two grids which will both be used below. Note that the Chebyshev-like grid has a high density of points near both endpoints whereas the uniform grid has an equal separation between every point and its nearest neighbors.

Because of Chebyshev’s identity $T_n(\cos(t)) = \cos(nt)$, this problem is equivalent to a Fourier extension problem of the form

$$g(t) \equiv f(\cos(t)) \approx \sum_{n=0}^{N} a_n \cos(nt), \quad t \in [0, \pi]$$

(29)

where the coefficients are determined by interpolation at points at

$$t_j = \arccos(x_j), \quad j = 1, 2, \ldots(N + 1), \quad t_j \in [0, \arccos(-\chi)]$$

(30)

Figure 4 shows that Chebyshev extension is convergent and successful even for the famous Runge function, $f(x) = 1/(1 + 25x^2)$, which has poles on the imaginary axis close to the expansion interval. There are however, some thorns among the roses.

The first one, well known from Fourier Extension, is that the interpolation matrix is very poorly conditioned as illustrated in Figure 5. The classic Chebyshev interpolation matrix, which applies when $\chi = 1$ and the interpolation interval is therefore the entire canonical Chebyshev interval, $x \in [-1, 1]$, has a condition number of $O(10)$. Unfortunately, the matrix condition number $\kappa$ grows exponentially as the domain parameter $\chi$ decreases. For fixed $\chi < 1$, the condition number also grows exponentially with increasing $N$, the matrix dimension. In the large blank area of the contour plot, the condition number is larger than $10^{16}$, which means that in usual Matlab precision where machine epsilon is $2.2 \times 10^{-16}$, we cannot guarantee that the matrix problem is solved with any accuracy at all. Fortunately, the matrix condition number, which is the ratio of the largest singular value divided by the smallest singular value, is a worst-case analysis that presumes that the inhomogeneous term in the matrix equation, which is the vector of samples of $f(x)$, projects entirely onto the mode of smallest singular value. The previous figure shows that in fact it is possible to obtain quite accurate approximations in spite of the ill conditioning for this example.

The second thorn is that the simplest grid is a uniform grid; constructing a Chebyshev-like grid for a domain with a convoluted boundary in two or three dimensions is a daunting task. Unfortunately, polynomial interpolation on a uniform grid is subject to the divergence known as the Runge Phenomenon. Fig. 6 is identical to Fig. 4 in every way except that the grid was changed from the Chebyshev-like grid to a lattice with an even spacing. Interpolation on the uniform grid diverges almost everywhere in the $\chi - N$ plane.

Runge showed that this divergence is due to the singularities of the target function. (For $f(x) = 1/(1 + 25x^2)$, the divergence-inducing singularities are simple poles at $x = \pm i/5$.) The function $f(x) = \cos(10x + \pi/4)$ is an entire function with no singularities except at infinity, and Fig. 7 confirms that for
some functions at least, accurate approximations on a uniform grid are indeed possible.

There are two important conclusions from these one-dimensional experiments:

1. Chebyshev extension can be successful with the right grid, at least in one dimension.
2. The uniform grid frequently fails with polynomial interpolation. There are three remedies discussed here in and in Parts I and III: boundary-clustered ["Chebyshev-like"] grids, hyperinterpolation, and radial basis functions.

7. Two-dimensional Extension: Tensor-Product Chebyshev Series on a Disk

The natural domain of a tensor product of Chebyshev polynomials is the unit square. Our goal is to apply them to the interior of a domain bounded by a squircle or other plane curve. Is it possible to extend the Chebyshev approximation from the square to the most extreme form of a squircle-bounded domain, the unit disk, as illustrated in Fig. 8? If all the grid points are confined to the disk, is it still possible for Chebyshev interpolation to yield useful results, at least within the disk? This section applies this ultimate approximation-on-the-disk stress test to the tensor product Chebyshev basis.

Let $D$ denote the maximum total degree. Let $mv(i)$ and $nv(i)$ denote vectors which map a single index $i$ to the degrees of individual polynomials in $x$ and $y$. (The are generated easily by the pseudocode $i=0; \text{for} m=0:mmax, \text{if } (2m+2n) \leq D \text{ & } m \leq n, i=i+1, mv(i)=m, nv(i)=n, \text{end, end, N=i.}$) Exploiting the full symmetry of the dihedral group $D_4$,

$$f_D(x,y) = \sum_{i=0}^{N} a(i) \{T_{2m(i)}(x)T_{2n(i)}(y) + T_{2n(i)}(x)T_{2m(i)}(y)\}$$

Here $i[m,n]$ is a matrix that maps the degree indices $(m,n)$ to a single index. The grid points are the usual Fourier-Chebyshev grid, restricted to one-eighth of the disk:

$$r_j = \cos \left( \frac{\pi}{2N_r} (j-1) \right), j = 1, \ldots N_r; \theta_k = \frac{\pi}{N_r} (2k-1), k = 1, 2, \ldots N_r/2$$

The plots below are a particularly stern test because $f(x,y) = 1/(2-x^2-y^2)$ is singular at the corner of the original square, that is, within the canonical domain for a tensor product Chebyshev basis. Nevertheless, Fig 10 shows that, provided the number of samples $P$ is somewhat larger than $N$, the number of basis functions, the approximation converges exponentially fast as the total degree $D \rightarrow \infty$. ($N = (D+1)^2$ when $D$ is even.) The contours become vertical for $P/N > 2$ to show that increasing this ratio beyond two does not improve accuracy. Near-interpolation with $P \approx N$ is a disaster.
8. Two-dimensional Extension: Zernike Polynomials for the Unit Square

The other extreme is to apply Zernike polynomials, which are a natural basis for the disk, to approximate functions on the square. As before, we exploit the $D_4$ symmetry of the square. For our examples, we choose $f(x, y)$ which are invariant under all operations of the eight-element group. The grid is the usual tensor product Chebyshev-Gauss grid, restricted to 1/8 of the square. The points are therefore

$$(x_i, y_j) \text{ such that } y_j < x_i, \quad x_i = \cos \left( \frac{2i - 1}{4M} \right), \quad y_j = \cos \left( \frac{2j - 1}{4M} \right),$$

where $M$ is a user-choosable integer, the cardinality of the one-dimensional grids. Fig. 11 is a visualization of such a grid.

Fig. 12 shows that hyperinterpolation errors are similar to those for Chebyshev polynomials on the disk. “Hypointerpolation” ($P < N$) and standard interpolation ($P \approx N$) are both uselessly inaccurate. However, the hyperinterpolation error falls exponentially fast as $P$ increases, leveling off when $P \approx 2N$, that is, least squares approximation with roughly twice as many points as basis functions.

The standard “triangular truncation” of the Zernike basis, which is $|m| \leq D$ accompanied by the restriction inherent in the definition of the Zernike polynomials, $|m| \leq n$, is a total degree truncation: all bivariate polynomials of total degree less than or equal to $D$ are included, but not polynomials of higher total degree.


Both Chebyshev and Zernike polynomials were applied to least squares hyperinterpolation on a domain bounded by a quadrifolium perturbed by a circle. The boundary curve is defined implicitly by

$$B(x, y) = \delta (x^2 + y^2) - ((x^2 + y^2)^3 - (x^2 - y^2)^2) = 0$$

where $\delta$ is a free parameter.

The grids were truncated uniform grids, the same for both basis sets, as illustrated for a typical case in Fig. 13.

Table 3 shows there is no significant difference between the Chebyshev and Zernike basis sets. These experiments demonstrate, fully consistent with the much more extensive experiments in Part III, that on irregular domains, high accuracy requires about twice as many points as basis points. This creates the odd situation that for a fixed number $P$ of interpolation points, accuracy is greatly improved by reducing the number of basis functions, but only until $P/N$ is larger than two.
Table 3: Hyperinterpolation on the quadrifolium-circle for $\delta = 1/1000$

The approximation target is $f(x, y) = 1/(x^2 + y^2 - 2)$. The grid is truncation of a uniform grid that covered one-fourth of the unit square; the size of the untruncated grid is also listed. All calculations used 16 decimal digit precision. Errors are maximum pointwise errors, that is, errors in the $L_\infty$ norm.

<table>
<thead>
<tr>
<th># points</th>
<th># basis funcs.</th>
<th>total degree</th>
<th>uniform grid</th>
<th>Chebyshev error</th>
<th>Zernike error</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>36</td>
<td>16</td>
<td>16 x 16</td>
<td>0.00336</td>
<td>5.57</td>
</tr>
<tr>
<td>52</td>
<td>25</td>
<td>16</td>
<td>16 x 16</td>
<td>1.58E-6</td>
<td>1.58E-6</td>
</tr>
<tr>
<td>115</td>
<td>100</td>
<td>36</td>
<td>24 x 24</td>
<td>0.980E-8</td>
<td>0.622E-4</td>
</tr>
<tr>
<td>115</td>
<td>81</td>
<td>32</td>
<td>24 x 24</td>
<td>0.151E-8</td>
<td>0.153E-5</td>
</tr>
<tr>
<td>115</td>
<td>64</td>
<td>28</td>
<td>24 x 24</td>
<td>0.624E-9</td>
<td>0.170E-5</td>
</tr>
<tr>
<td>115</td>
<td>49</td>
<td>24</td>
<td>24 x 24</td>
<td>0.505E-8</td>
<td>0.505E-8</td>
</tr>
<tr>
<td>204</td>
<td>225</td>
<td>56</td>
<td>32 x 32</td>
<td>0.224E-2</td>
<td>0.288E-3</td>
</tr>
<tr>
<td>204</td>
<td>169</td>
<td>48</td>
<td>32 x 32</td>
<td>0.115E-7</td>
<td>0.245E-5</td>
</tr>
<tr>
<td>204</td>
<td>121</td>
<td>40</td>
<td>32 x 32</td>
<td>0.255E-10</td>
<td>0.785E-7</td>
</tr>
<tr>
<td>204</td>
<td>81</td>
<td>32</td>
<td>32 x 32</td>
<td>0.173E-10</td>
<td>0.173E-10</td>
</tr>
<tr>
<td>458</td>
<td>441</td>
<td>80</td>
<td>48 x 48</td>
<td>0.0968</td>
<td>0.1950</td>
</tr>
<tr>
<td>458</td>
<td>361</td>
<td>72</td>
<td>48 x 48</td>
<td>0.107E-3</td>
<td>0.232E-3</td>
</tr>
<tr>
<td>458</td>
<td>289</td>
<td>64</td>
<td>48 x 48</td>
<td>0.735e-7</td>
<td>0.336E-6</td>
</tr>
<tr>
<td>458</td>
<td>225</td>
<td>56</td>
<td>48 x 48</td>
<td>0.391e-9</td>
<td>0.110E-8</td>
</tr>
<tr>
<td>458</td>
<td>169</td>
<td>48</td>
<td>48 x 48</td>
<td>0.455e-11</td>
<td>0.948E-11</td>
</tr>
</tbody>
</table>

To be sure, the Chebyshev basis converges much rapidly for a fixed number of points $P$ to the best accuracy obtainable for that $P$. However, the parametric range where Chebyshev is superior to Zernike generates inaccurate and inefficient approximations. For a given $P$, it is cheapest and most accurate to choose $P/N$ is in the range two to three, and there the Chebyshev/Zernike differences melt away to nothing.

We did similar experiments for the same two basis sets for a different $f(x, y)$ and a different domain. The function to be approximated is

$$f(x, y) = \cos(\pi x^4) \cos(\pi y^4)/(x^2 + y^2 - 2)$$  \hfill (34)

The boundary is the squircle defined implicitly by

$$B(x, y) = 1 - (x^4 + y^4) - 1 = 0$$  \hfill (35)

Results are shown in Table 4. Hyperinterpolation on the squircle proved so similar to similar approximations on the quadrifolium that, besides offering this table without comment, no further discussion is unwarranted.
Table 4: Hyperinterpolation of $f(x, y) = \cos(\pi x^4) \cos(\pi y^4)/(x^2 + y^2 - 2)$ in a domain bounded by a squircle.

The grid was truncated from the uniform grid on a quarter of the unit square; the size of the untruncated grid is also listed. All calculations used 16 decimal digit precision. Errors are maximum pointwise errors, that is, errors in the $L_\infty$ norm.

<table>
<thead>
<tr>
<th># points</th>
<th># basis funcs.</th>
<th>total degree</th>
<th>uniform grid</th>
<th>Chebyshev error</th>
<th>Zernike error</th>
</tr>
</thead>
<tbody>
<tr>
<td>66</td>
<td>9</td>
<td>8</td>
<td>12 × 12</td>
<td>0.202</td>
<td>0.202</td>
</tr>
<tr>
<td>66</td>
<td>16</td>
<td>12</td>
<td>12 × 12</td>
<td>0.192</td>
<td>0.193</td>
</tr>
<tr>
<td>66</td>
<td>25</td>
<td>16</td>
<td>12 × 12</td>
<td>0.0471</td>
<td>0.047</td>
</tr>
<tr>
<td>66</td>
<td>36</td>
<td>20</td>
<td>12 × 12</td>
<td>0.0116</td>
<td>0.0116</td>
</tr>
<tr>
<td>66</td>
<td>49</td>
<td>24</td>
<td>12 × 12</td>
<td>0.346</td>
<td>2.4E5</td>
</tr>
<tr>
<td>66</td>
<td>56</td>
<td>26</td>
<td>12 × 12</td>
<td>3.5E8</td>
<td>2.4E5</td>
</tr>
<tr>
<td>66</td>
<td>64</td>
<td>28</td>
<td>12 × 12</td>
<td>7.5E7</td>
<td>2.9E</td>
</tr>
</tbody>
</table>

| 267      | 25             | 16           | 24 × 24      | 0.0211          | 0.0211        |
| 267      | 36             | 20           | 24 × 24      | 0.0167          | 0.00168       |
| 267      | 49             | 24           | 24 × 24      | 0.00108         | 0.021         |
| 267      | 64             | 28           | 24 × 24      | 0.000151        | 0.000151      |
| 267      | 81             | 32           | 24 × 24      | 2.00E-5         | 2.0E-5        |
| 267      | 100            | 36           | 24 × 24      | 5.32E-6         | 6.67E-6       |
| 267      | 121            | 40           | 24 × 24      | 8.06E-6         | 1.21E-5       |
| 267      | 144            | 44           | 24 × 24      | 3.11E-5         | 0.00575       |
| 267      | 169            | 48           | 24 × 24      | 1.67            | 0.156         |
| 267      | 196            | 52           | 24 × 24      | 2.33            | 4.22          |
| 267      | 225            | 56           | 24 × 24      | 2.19            | 1.26E3        |

| 417      | 49             | 24           | 30 × 30      | 4.84E-4         | 4.84E-4       |
| 417      | 64             | 30           | 30 × 30      | 3.75E-5         | 5.46E-5       |
| 417      | 121            | 42           | 30 × 30      | 1.65E-7         | 9.05E-7       |
| 417      | 169            | 48           | 30 × 30      | 4.15E-8         | 7.75E-4       |
| 417      | 196            | 54           | 30 × 30      | 8.16E-4         | 2.84E-3       |
| 417      | 256            | 60           | 30 × 30      | 0.502           | 0.381         |
| 417      | 361            | 72           | 30 × 30      | 0.702           | 2.60E2        |
10. Zernike versus Chebyshev: A Draw

Although the Zernike and tensor product Chebyshev basis sets are, with
total degree truncation and identical grids, exactly the same in infinite precision,
spectral extension is ill-conditioned. It therefore seemed logical that polynomials
which oscillate in radial patterns would be numerically inferior for the square,
and that tensor products of Chebyshev polynomials, with nodal surfaces parallel
to the boundaries, would be superior in finite precision for the square.

Exhaustive numerical experiments failed to find much evidence of a significant
difference between Zernike and Chebyshev polynomials. As one final but
representative example, we applied both Zernike and Chebyshev polynomials to
the same problem on the same tensor product Chebyshev grid on the square.
Fig. 14 shows that the differences between the two are not significant.

11. Summary

In infinite precision arithmetic, all polynomial bases with the same truncation-
by-total-degree are identical. In finite precision arithmetic, it is well known that
the monomial basis is ill-conditioned with the condition number growing expo-
nentially with increasing \( N \). Zernike polynomials are a natural basis for the
disk; tensor products of Chebyshev polynomials oscillate uniformly within the
unit square and explode exponentially outside it. We expected that there would
be huge numerical differences between the two. To our surprise, there was lit-
tle difference. We show that Chebyshev hyperinterpolation with total degree
basis truncation converges at a geometric, exponential rate to extremely high
accuracy even in the Chebyshev- unfavor able extreme of the unit disk. Similarly,
Zernike hyperinterpolation is triumphant even for the undisk-like extreme of the
square.

The Stone-Weierstrass Theorem asserts that a smooth function can always
be approximated by a polynomial even in a convoluted domain. We have shown
that for a squircle-bounded domain, least squares hyperinterpolation with \( P \geq
2N \) yields a geometrically convergent approximation when \( f(x, y) \) is smooth.

Hyperinterpolation cannot guarantee convergence on a uniform grid unless
the number of interpolation points \( P \) is \( O(N^2) \) [32, 13]. The ideal, safe grid is
very non-uniform with a high density of points along the boundaries, mimicking
the one-dimensional Chebyshev grid. Experimentally, however, a much simpler
construction worked most of the time: the “truncated uniform” grid. This
is built by first embedding the domain in a square, constructing a uniformly-
spaced, tensor product grid for the square, and then deleting all points that lie
outside the domain \( \Omega \). If the sign of the boundary function is normalized so
that \( B \) is negative on the interior, it is then trivial to test the inclusion of a
point merely by examining the sign of \( B \) at that point.

As noted earlier by Bossavit [5, 6] (and others he cites), an arbitrary function
on a domain invariant under the eight operations of the \( D_4 \) symmetry group can
be decomposed into six parts. The interpolation problem can then be reduced
from an \( N \times N \) matrix equation to four matrix problems of size \( N/8 \) and two
of size $N/4$ at a cost which is $3/64$ that of the original problem, assuming use of Gaussian elimination. (The two larger subproblems share the same interpolation matrix.) For simplicity, we here focused on the completely symmetric part, but everything we have done can be applied equally well to the other five subproblems.

Another frontier is to apply RBFs and polynomial interpolation to geometries with holes, complicated curvature and so on. With due respect to the nonconstructive theory of Weierstrass and Stone, it seems likely that very complicated domains will “break” RBFs and polynomial interpolation, unless memory and runtime requirements beyond the capabilities of a desktop machine are deemed acceptable. However, we are a long way from knowing what these limits are.

Acknowledgments: This work was supported by the National Science Foundation through grants OCE 0451951 and OCE 1059703. Shan Li was partially supported by NSFC grant 11171209 of China. We thank Jianping Xiao for his careful proofreading.

Appendix A. Uniform Hexagonal Grid

Fig. A.15 shows a hexagonal grid embedded in the unit square. Each hexagon whose side is of length $s$ can be decomposed into six equilateral triangles with sides also of length $s$, so this type of grid is also labeled as an “equilateral triangular” grid. Fornberg, Flyer and Russell tout the advantages of hexagonal grids for RBFs and finite differences [22]. Each point is surrounded by six nearest neighbors instead of the four of a rectangular grid. The condition number of the interpolation matrix is significantly improved [12]. The accuracy of spatial differentiation is also markedly better.

Unfortunately, a hexagonal grid has severe disadvantages for a squircle-bounded domain. First, the vertical separation between grid rows is smaller by a factor of $\sqrt{3}/\sqrt{2} \approx 0.866$ than the vertical rows. (This is a consequence of the geometry of an equilateral triangle, six of which fit together to make the hexagon, as illustrated in Fig. A.16.) One could of course stretch the grid in the vertical by the reciprocal of this factor, but this destroys the symmetry of the grid, reduces accuracy and increases ill conditioning. Thus, the hexagonal grid does not pack tightly inside the unit square, which is obviously a disadvantage in the limit that the squircle tends to the square, and likely for finite squircle exponent $\nu$ as well. The unequal horizontal and vertical spacings is a generic drawback of hexagonal grids, inconvenient for general domains.

Another serious flaw is narrow in the sense that it arises only for domains of $D_4$ group invariance such as squircle-bounded and perturbed-quadrifolium-bounded regions. The flaw for $D_4$-invariant domains is that although the hexagonal grid is invariant with respect to reflections about the coordinate axes, the diagonal dotted line in Fig. A.15 makes it obvious that the grid is not invariant with respect to reflections about this diagonal symmetry axis. This means that it is not possible to fully exploit the symmetry of a squircle-bounded domain using a hexagonal grid.
Because of these drawbacks, we have not used the hexagonal grid here.

References


Figure 2: The Zernike polynomials $Z_{10}^0(x, y)$, $Z_{20}^0(x, y)$ and $Z_{40}^0(x, y)$, which are functions only of $r = \sqrt{x^2 + y^2}$, plotted versus $r$. These polynomials, and the non-radially symmetric Zernike polynomials not illustrated, oscillate when $r < 1$, but grow exponentially for $r > 1$.

Figure 3: A comparison of the Chebyshev-like and uniform grids on the interval $x \in [-\chi, 1]$ where $\chi = 1/2$. 
Figure 4: Contours of the base 10 logarithm of the maximum pointwise error space (error in the $L_{\infty}$ norm) for the approximation of $f(x) = 1/(1 + 25x^2)$ by an $N$-term Chebyshev series when all the interpolation points are confined to the interval $x \in [-\chi, 1]$. The grid is Chebyshev-like in that the points are separated by a distance of $O(1/N^2)$ near both endpoints of the interval $x \in [\chi, 1]$. 
Figure 5: Same but showing the contours of the base 10 logarithm of the condition number of the interpolation matrix.
Figure 6: Similar to Fig. 4 except that the grid is spatially uniform, $x_j = x_j^{\text{uniform}} = -\chi + (1 + \chi)j/N, j = 1, 2, \ldots, N$. Contours of the base 10 logarithm of the maximum pointwise error (error in the $L_\infty$ norm) for the approximation of $f(x) = 1/(1 + 25x^2)$ by an $N$-term Chebyshev series when all the interpolation points are confined to the interval $x \in [-\chi, 1]$. 
Figure 7: Same as the previous figure except \( f(x) = \cos(10x + \pi/4) \). Errors grow for larger \( N \) (instead of the hoped-for decrease) because uniform grid polynomial interpolation is very ill-conditioned. The apparent divergence for large and increasing \( N \) for this \( f(x) \) can be removed by multiple precision arithmetic. The Runge divergence in the preceding figure, in contrast, cannot be converted to convergence even by arbitrarily high precision computations.
Figure 8: Schematic of Chebyshev tensor product extension. Although the basis is canonical for the unit square, all grid points (and errors of the approximation) are restricted to the unit disk (shaded). The error norm is the maximum pointwise error, restricted to points on the unit disks. The approximated function, \( f(x, y) = 1/(2 - x^2 - y^2) \), is nasty because it is singular everywhere on the circle of radius \( \sqrt{2} \) [dashed circle], which includes the corners of the unit square.
Figure 9: Contours of the base-10 logarithm of the $L_{\infty}$ error norm for approximation of $f(x,y) = 1/(2 - x^2 - y^2)$ on the unit disk by a tensor product Chebyshev basis, symmetrized to be invariant under the full $D_4$ group with a total degree truncation. The highest retained basis functions are $T_{50}(x) + T_{50}(y)$ and $T_{48}(x)T_2(y) + T_2(x)T_{48}(y)$, $T_{46}(x)T_4(y) + T_4(x)T_{46}(y)$ and so on. The vertical axis is the ratio $P/N$ [points/number of basis functions]; the error “saturates” when $P/N$ is larger than two.
Figure 10: Same case as previous figure, but showing contours of the base-10 logarithms of the condition number of the Vandermonde matrix for approximation of \( f(x, y) = 1/(2 - x^2 - y^2) \) on the unit disk by a tensor product Chebyshev basis, symmetrized to be invariant under the full \( D_4 \) group with a total degree truncation.
Figure 11: A 231 point tensor product Chebyshev grid in 1/8 of the square as used for Zernike-on-the-square experiments. ($M = 22$.)
Figure 12: $L_\infty$ norm of the error for the approximation of $f(x,y) = \sin(8(x^4 + y^4)) \cos(8(x^2+y^2))/(2 + x^2 + y^2)$ by Zernike polynomials $Z_{mn}^m(r/\sqrt{2}, \theta)$ by hyperinterpolation on the unit square using a tensor product Chebyshev grid in 16 decimal digit precision. Each of three curves is the error for a fixed number of basis functions $N$ and variable numbers $P$ of hyperinterpolation points. Hypointerpolation ($P < N$) always gave large errors, so each curve begins with $P \approx N$. 

$Zernike hyperinterpolation on the unit square$

$f=\sin(8(x^4 + y^4)) \cos(8(x^2+y^2))/(2 + x^2 + y^2)$
Figure 13: Truncated uniform grid for the perturbed-quadrifolium for $\delta = 1/1000$. Because the domain is invariant under the eight operations of the $D_4$ symmetry group, it is sufficient use grid points only in one eighth of the domain bounded by the perturbed-quadrifolium as explained in our previous work. Similarly, the basis functions were symmetrized to be invariant with respect to the same symmetry group.
Figure 14: $L_{\infty}$ norm of the error for the approximation of $\sin(8(x^4 + y^4)) \cos(8(x^2 + y^2))/(2 + x^2 + y^2)$ by Zernike polynomials $Z_m^n(r/\sqrt{2}, \theta)$ and also by tensor product Chebyshev polynomials by hyperinterpolation on the unit square in 16 decimal digit precision. The Chebyshev points were generated by taking a tensor product of the one-dimensional grids generated by $x_j = \cos(\pi(2j - 1)/(4M)), j = 1, 2, \ldots M$ and deleting all points outside the octant $x \leq y, y > 0$, giving $P = M(M - 1)/2$ points. The number of polynomials of total degree $D$ which are fully symmetric under the $D_4$ group is $N = ((D/4)+1)^2$. Every Zernike polynomial in this symmetry class has a total degree $D$ which is a multiple of four; to ensure that the number of points slightly exceeded twice the number of basis functions, we set $D = 4j$ and $M = 2j + 3$ where $j = 1, 2, \ldots$. The reason we choose a ratio $P/N$ slightly large than two is that hyperinterpolation here and on the sphere is usually inaccurate until $P/N \geq 2$. 

Zernike

 tensor Chebyshev
Figure A.15: A hexagonal grid (black disks) and the hexagons that define the grid. The dashed lines are the sides of the unit square. If $s$ denotes one side of a hexagon, then the horizontal grid spacing is $s$, as is marked by a line segment underneath the grid, but the vertical spacing is $(1/2)\sqrt{3}s = 0.866s$, the vertex-to-midpoint length of the perpendicular bisector of one side of the equilateral triangle of length $s$, as indicated by the line segment to the right of the grid. The dotted line is the diagonal $y = x$; the squircle is invariant to reflection about this line, but the hexagonal grid obviously is not.
Figure A.16: An equilateral triangle, one of the building blocks of the hexagonal grid. Each vertex is a grid point. Each equilateral triangle can be subdivided into two 30-60-90 right triangles, so named because of the angles in degrees at each vertex as labeled. The horizontal separation is the length of the side of the triangle, \( s \). The vertical separation between grid points is the perpendicular distance from the upper vertex to the to midpoint of the base. This distance is not \( s \), but only 0.866\( s \).