Orthogonality and Computation

D. M. Topa[1,2,3] **and P. F. Embid**[1,2]

[1]Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM, USA
[2]Los Alamos National Laboratory, Los Alamos, NM, USA
[3]Engility Corporation, U.S.A.C.E. Engineer Research and Development Center,
Information Technology Laboratory, Vicksburg MS, USA

Abstract— *The property of orthogonality is predicated upon the specifications of a domain and a topology. The orthogonality of the continuum is violated in the computational domain as evidenced by poor convergence and numerical oscillations. Penalties are significant numerical errors and a substantial increase in computation time. By using linear independence, exact solutions are found in specific instances.*

Keywords: orthogonality, linear independence, Hilbert space, Lebesgue integration, domain topology

1. Introduction

Orthogonality and projection are two facets of the same gem. They are foundation concepts in many areas of science and engineering. For example, the mathematics of quantum mechanics and quantum field theory are the embodiment of the power of orthogonal projection. The least squares method delivers an orthogonal projection. A great deal of mathematics is based upon orthogonality. Computations are too often corrupted by the inappropriate presumption of orthogonality. This paper explores how this malady distorts numerical computations and presents a more suitable schema.

Common knowledge is that the trigonometric functions $\sin nx$ and $\cos nx$ (and their linear combination e^{inx}) are orthogonal over both the continuous domain $\Omega = \{x\colon -\pi \le x \le \pi\}$ and the uniformly sampled discrete domain $\sigma = \left\{x\colon x_\nu = -\pi + \frac{\nu}{n}\pi\right\}_{\nu=0}^{2n}$. A branch of mathematics built on these principles is Fourier analysis, which approximates arbitrary functions as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Uncommon knowledge is that the sine and the cosine are the *only* functions orthogonal in *both* the continuum and in discrete space. This explains why there is no such thing as a discrete Legendre transform or a fast discrete Zernike transform.

Part of the confusion may stem from the famous case, Parseval's identity, where the two spaces are connected. The relation between smoothness in physical space and decay of the Fourier amplitudes is so fundamental in mathematics

that it gives rise to the so-called Sobolev spaces, where the smoothness of the function $f(x)$ is understood in terms of the L^2−norm, and the corresponding decay of the coefficients is given in the related l^2−norm via

$$\int_{-\pi}^{\pi} \left| \frac{d^k f}{dx^k} \right|^2 dx = \sum_{n=-\infty}^{n=\infty} n^{2k} |c_n|^2 . \tag{1.1}$$

The spaces L^2 and l^2 are distinct and, excluding sine and cosine, there are no functions which are orthogonal in both.

2. Spaces and Topologies

Troubles arise when moving from the continuous space L^2 to the discrete space of l^2. This corresponds to moving from the continuum, the theoretical realm of the chalkboard, to discrete space, the realm of computer calculation. Either measurement or computation imply a discrete topology which sacrifices orthogonality.

2.1 Continuous vs. discrete

The stage is set with definitions of domains, norms, and membership for both topologies. Start with Ω, a continuous interval on the real number line. Given the boundary points, $[a, b]$ where $a < b$ this domain is formally defined as

$$\Omega = \{x \in \mathbb{R}\colon a \le x \le b\}. \tag{2.1}$$

For the discrete case there is but a finite sequence of μ points

$$\sigma = \{x_1, x_2, \ldots, x_\mu\} = \{x_\nu\}_{\nu=1}^{\mu} \tag{2.2}$$

ordered such that

$$a \le x_1 < x_2 < \cdots < x_\mu \le b. \tag{2.3}$$

Such a set represents a partition of the continuous interval Ω. The natural norms are introduced to measure distance. For continuous topologies

$$\|F(x)\|_{L^2}^2 = \int_{\Omega} F(x)\overline{F(x)}dx. \tag{2.4}$$

Integration is in the Lebesgue sense. For discrete topologies

$$\|f(x)\|_{l^2}^2 = \sum_{x \in \sigma} f(x)\overline{f(x)}. \tag{2.5}$$

The set σ may form a uniform mesh where $x_{k+1} - x_k = \Delta$, $k = 1, 2, \ldots, \mu-1$. Reserve capital letters for the continuum, e.g., L^2, and lower case letters for discrete spaces, e.g., l^2.

The collection of all functions which are square integrable in the continuum is $L^2(\Omega)$. A function $F(x)\colon \mathbb{R} \mapsto \mathbb{C}$ is an element in this space iff the norm is finite,

$$F(x) \in L^2(\Omega) \quad \Longleftrightarrow \quad \int_\Omega F(x)\overline{F(x)}dx < \infty. \quad (2.6)$$

The collection of all functions which are square summable is $l^2(\sigma)$. A function $f(x)\colon \mathbb{R} \mapsto \mathbb{C}$ is an element in this space iff the norm is finite,

$$f(x) \in l^2(\sigma) \quad \Longleftrightarrow \quad \sum_{x \in \sigma} f(x)\overline{f(x)} < \infty. \quad (2.7)$$

With the stage set, the task at hand is resolving a target function in the bases of a complete and linearly independent set of functions over the domain. Let the highest order in the expansion be d for degree of fit, and for the continuum case where $x \in \Omega$ the target function is F, approximated as

$$F(x) \approx a_0 G_0(x) + a_1 G_1(x) + \cdots + a_d G_d(x). \quad (2.8)$$

In discrete space $x \in \sigma$ the target function f is approximated as

$$f(x) \approx b_0 g_0(x) + b_1 g_1(x) + \cdots + b_d g_d(x). \quad (2.9)$$

2.2 Riesz–Fischer theorem

The Riesz–Fischer theorem [2, p. 330] is a powerful result about the convergence of Cauchy sequences in L^2 spaces:[1]

Theorem 1 (Riesz–Fischer): Let $\{\phi_n\}$ be an orthonormal sequence of functions on Ω and suppose $\sum |a_n|^2$ converges. Denote the partial sum as

$$s_d = a_0 \phi_0 + a_1 \phi_1 + \cdots + a_d \phi_d.$$

There exists a function $F \in L^2(\Omega)$ such that $\{s_d\}$ converges to F in $L^2(\Omega)$, and such that

$$F = \sum_{k=0}^{\infty} a_k \phi_k, \quad (2.10)$$

almost everywhere.

Proof: The proof is both a staple in books on functional analysis and outside the scope of this article. ∎

This existence theorem motivates a search for the amplitudes a_k. Section §4.3.2 exhibits the problems that arise when the basis set of functions is not orthogonal.

3. Background

This section provides the rudiments for finding a least squares solution for the amplitudes and discusses the benefits of orthogonality and linear independence. (Insightful background can be found in [3, §4.6, §5.13], [4, ch 8].)

[1]The theorem generalizes to $1 \leq p < \infty$.

3.1 Least squares in L^2

Given a domain such as Ω in (2.1), a function $F(x) \in L^2(\Omega)$, and a sequence of basis functions $\{G_k\}_{k=0}^d \in L^2(\Omega)$, the least squares solution a_{LS} in the continuum is cast as

$$a_{LS} = \Big\{ a \in \mathbb{C}^{d+1} \colon$$
$$\int_\Omega \Big| F(x) - \sum_{k=0}^d a_k G_k(x) \Big|^2 \, dx \text{ is minimized} \Big\}. \quad (3.1)$$

3.1.1 Linear independence

If the sequence of functions $\{G_k\}$ is *linearly independent* the resulting linear system,

$$\mathbf{A}^*\mathbf{A}a = \mathbf{A}^*F, \quad (3.2)$$

has full rank and is not singular. A simpler tool such as the normal equations can be used to find the solution vector a. The matrix form of these normal equations is

$$\underbrace{\begin{bmatrix} \langle G_0|G_0\rangle & \langle G_0|G_1\rangle & \cdots & \langle G_0|G_d\rangle \\ \langle G_1|G_0\rangle & \langle G_1|G_1\rangle & \cdots & \langle G_1|G_d\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle G_d|G_0\rangle & \langle G_d|G_1\rangle & \cdots & \langle G_d|G_d\rangle \end{bmatrix}}_{\mathbf{A}^*\mathbf{A}} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} \langle F|G_0\rangle \\ \langle F|G_1\rangle \\ \vdots \\ \langle F|G_d\rangle \end{bmatrix}}_{\mathbf{A}^*F}. \quad (3.3)$$

The inner product defines the matrix elements as

$$\langle G_j|G_k\rangle_{L^2} = \int_\Omega G_j(x)\overline{G_k(x)}dx. \quad (3.4)$$

The product matrix $\mathbf{A}^*\mathbf{A}$ is symmetric positive definite, and the eigenvalues are non-negative. The least squares solution is

$$a_{LS} = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*F \quad (3.5)$$

3.1.2 Orthogonality

The Reisz-Fischer theorem motivates the use of an *orthogonal set* of basis functions. In this case, the inner products are simplified to

$$\langle G_j|G_k\rangle = \xi_j \delta_{jk}, \quad (3.6)$$

where δ_{jk} is the Kronecker delta function. Now the linear system decouples and the amplitudes can be solved mode-by-mode.

$$\begin{bmatrix} \xi_0 & 0 & \cdots & 0 \\ 0 & \xi_1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \xi_d \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{bmatrix} = \begin{bmatrix} \langle F|G_0\rangle \\ \langle F|G_1\rangle \\ \vdots \\ \langle F|G_d\rangle \end{bmatrix}. \quad (3.7)$$

The amplitudes are computed independently of one another

$$a_k = \frac{\langle F|G_k\rangle_{L^2}}{\langle G_k|G_k\rangle_{L^2}} = \frac{\langle F|G_k\rangle_{L^2}}{\xi_k}, \quad k = 0, 1, 2, \dots, d. \tag{3.8}$$

For clarity, the inner products are marked to remind us of the host space topology.

A great benefit of decoupled systems is the amplitudes do not change as the order of fit is increased; given $d_1 \leq d_2$, $\{a_k\}_{k=0}^{d_1} \subseteq \{a_k\}_{k=0}^{d_2}$.

3.2 Least squares in l^2

Given a domain such as σ in (2.2), a function $f(x) \in l^2(\sigma)$, and a sequence of basis functions $\{g_k\}_{k=0}^d \in l^2(\sigma)$, the least squares solution b_{LS} in the discrete space is written as

$$b_{LS} = \Big\{ b \in \mathbb{C}^{d+1} :$$
$$\sum_{\nu=1}^{\mu} \left(f(x_\nu) - \sum_{k=0}^d b_k g_k(x_\nu) \right)^2 \text{ is minimized} \Big\}. \tag{3.9}$$

as seen in (3.1). Different capitalization distinguishes the basis functions in $G_k(x) \in L^2\Omega$ from the basis functions $g_k(x) \in l^2(\sigma)$. For orthogonal functions, the *only* time $G_k(x) = g_k(x)$ is when they are sines and cosines.

3.2.1 Linearity

When the function sequence $\{g_k\}_{k=0}^d$ is linearly independent, the normal equations create a full rank linear system as seen in 3.3

$$\begin{bmatrix} \langle g_0|g_0\rangle & \langle g_0|g_1\rangle & \cdots & \langle g_0|g_d\rangle \\ \langle g_1|g_0\rangle & \langle g_1|g_1\rangle & \cdots & \langle g_1|g_d\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle g_d|g_0\rangle & \langle g_d|g_1\rangle & \cdots & \langle g_d|g_d\rangle \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_d \end{bmatrix}$$
$$= \begin{bmatrix} \langle f|g_0\rangle \\ \langle f|g_1\rangle \\ \vdots \\ \langle f|g_d\rangle \end{bmatrix}, \tag{3.10}$$

with the matrix elements as

$$\langle g_j|g_k\rangle_{l^2} = \sum_{x \in \sigma} g_j(x) g_k(x). \tag{3.11}$$

3.2.2 Orthogonality

In cases[2] where the basis functions are orthogonal, the linear system decouples, and amplitudes can be computed

---

[2]Of course sine and cosine are orthogonal; other functions can be constructed to be orthogonal over a discrete domain [1, §8.2].

mode-by-mode

$$\begin{bmatrix} \eta_0 & 0 & \cdots & 0 \\ 0 & \eta_1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \eta_d \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_d \end{bmatrix} = \begin{bmatrix} \langle f|g_0\rangle \\ \langle f|g_1\rangle \\ \vdots \\ \langle f|g_d\rangle \end{bmatrix}. \tag{3.12}$$

The solution set is

$$b_k = \frac{\langle f|g_k\rangle_{l^2}}{\langle g_k|g_k\rangle_{l^2}}, \quad k = 0, 1, 2, \dots, d. \tag{3.13}$$

In contrast to (3.4) the inner product here is

$$\langle f|g_k\rangle_{l^2} = \sum_{j=0}^d f(x_j) \overline{g_k(x_j)}. \tag{3.14}$$

The linear system of normal equations in (3.10) is effectively diagonalized.

3.3 Summary

An orthogonal set of basis functions presents a significant advantage. The linear system decouples and the amplitudes can be computed directly using either (3.7) or (3.12) depending upon the topology. An orthogonal basis is also a linearly independent basis. Without orthogonality, and with only linear independence, a large linear system such as (3.3) or (3.10) must be solved. Without orthogonality, and without linear independence, the linear systems lack full rank and the solution demands tools like **QR** decomposition or the singular value decomposition.

4. Demonstration

The choice of a polynomial basis is driven by the domain, the domain of choice is the continuous interval $\Omega = \{x \in \mathbb{R} : -1 \leq x \leq 1\}$. The following demonstration uses the Legendre polynomials to illustrate points of emphasis.

4.1 Definitions

These functions are monic polynomials, orthogonal over the continuous set $\Omega = \{x : -1 \leq x \leq 1\}$. As expected, they are not orthogonal over the discrete and well-ordered set $\sigma = \{x_k : -1 \leq x_k \leq 1\}$, $k = 1, 2, \dots, \mu$.

The polynomials can be defined as the set of functions which solve Legendre's differential equation,

$$\frac{d}{dx}\left((1-x^2)\frac{d}{dx}P_n(x) \right) + n(n+1)P_n(x) = 0. \tag{4.1}$$

A recipe for constructing the set starts with $P_0(x) = 1$ and $P_1(x) = x$ and employs the recursion relationship

$$P_{n+1}(x) = xP_n(x) - \frac{n^2}{4n^2 - 1}P_{n-1}(x). \tag{4.2}$$

This format demonstrates the definite parity of the polynomials, which toggles as the order increases in unit steps. The Legendre polynomials through order d are generated by

applying a Gram-Schmidt orthogonalization [3, p. 309] to the monomial functions $\left\{x^k\right\}_{k=0}^d$ through order d,

$$\mathcal{P}_d(x) = x^d - \sum_{k=0}^{d-1}\langle \mathcal{P}_k(x)|x^d\rangle. \qquad (4.3)$$

$\{\mathcal{P}_n(x)\}_{n=0}^{\infty}$ is a sequence of orthogonal functions while $\{P_n(x)\}_{n=0}^{\infty}$ is a sequence of orthogonal and monic functions; the difference is a set of scaling factors.

As evidenced in (4.2), the parity of the Legendre polynomials is the same as the parity of the index. For example, when n is an odd number $P_n(x)$ is an odd function,

$$P_n(-x) = \begin{cases} P_n(x) & n \text{ even} \\ -P_n(x) & n \text{ odd} \end{cases}. \qquad (4.4)$$

The norm is given as

$$\langle P_m|P_n\rangle = \frac{2}{2n+1}\delta_{mn}. \qquad (4.5)$$

4.2 Computational efficiency

The Legendre polynomials are orthogonal in the continuum, which leads to a trivial form for the amplitudes,

$$\begin{aligned} a_k &= \frac{\langle F|P_k\rangle_{L^2}}{\langle P_k|P_k\rangle_{L^2}} \\ &= \frac{2k+1}{2}\int_{-1}^{1} F(x)P_k(x)dx, \quad k = 0, 1, 2, \ldots, d, \end{aligned} \qquad (4.6)$$

as seen in (3.8). But in the discrete domain of computation, the off-diagonal matrix elements in (3.10) such as

$$\langle P_2(x)|P_4(x)\rangle = \sum_{\nu=1}^{\mu} P_2(x_\nu) P_4(x_\nu)dx \qquad (4.7)$$

are nonzero, demanding solution of the complete linear system. Denuded of orthogonality, the Legendre polynomial set is a cumbersome combination of monomials. Table 1 compares functional forms for monomials and Legendre polynomials.

Table 1: Functional forms for the monomials and the Legendre polynomials.

order	monomial	Legendre
0	1	1
1	x	x
2	x^2	$\frac{1}{2}\left(3x^2 - 1\right)$
3	x^3	$\frac{1}{2}\left(5x^3 - 3x\right)$
4	x^4	$\frac{1}{8}\left(35x^4 - 30x^2 + 3\right)$
5	x^5	$\frac{1}{8}\left(63x^5 - 70x^3 + 15x\right)$

Comparing the monomials to the Legendre polynomials order by order, the succinctness of the monomials is apparent.

4.2.1 Monomials

Showing that the monomials are a minimal spanning set requires establishing that they are linearly independent.

Theorem 2 (Linear independence of the monomials): The monomial set $\left\{1, x, x^2, \ldots, x^d\right\}$ with $d \geq 1$ is linearly independent over any continuous interval including the origin.

Proof: The crux of the proof uses induction to show that the Wronskian

$$W(x) = \begin{vmatrix} 1 & x & x^2 & \cdots & x^d \\ 0 & 1 & 2x & & dx^{d-1} \\ 0 & 0 & 2 & & d(d-1)x^{d-2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & d! \end{vmatrix} \qquad (4.8)$$

evaluates to $W(0) = \prod_{k=0}^d k! \neq 0$. ∎

4.2.2 Spans

Consider two different spans for the Legendre polynomials S_P, and for the monomials S_m,

$$S_P = \text{sp}\left\{\begin{bmatrix} P_0(x) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ P_1(x) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ P_d(x) \end{bmatrix}\right\},$$

$$S_m = \text{sp}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ x^d \end{bmatrix}\right\}. \qquad (4.9)$$

Both spans are linearly independent. In the limit $d \to \infty$, both spans are complete. The span S_P is orthogonal over Ω, neither span is orthogonal over σ.

Theorem 3 (Equivalence of spans): Let $\{S_m\}^d$ be the span of a Hilbert space described by the monomial set $\left\{1, x, x^2, \ldots, x^d\right\}$, and let $\{S_P\}^d$ be the span of a Hilbert space described by the Legendre polynomial set $\{P_0(x), P_1(x), P_2(x), \ldots, P_d(x)\}$. The spans $\{S_m\}^d$ and $\{S_P\}^d$ are equivalent.

Proof: Proof is based upon the construction process in (4.3). ∎

4.2.3 Transformations

Given the equivalence of the Hilbert spaces spanned by S_m and S_P, there exists an affine transformation to move between the Legendre basis and the monomial basis. One

such transformation matrix is shown in (4.10) for $d = 5$,

$$T_5 = \frac{1}{8} \begin{array}{c} \begin{array}{cccccc} P_0 & P_1 & P_2 & P_3 & P_4 & P_5 \end{array} \\ \begin{bmatrix} 8 & 0 & -4 & 0 & 3 & 0 \\ 0 & 8 & 0 & -12 & 0 & 15 \\ 0 & 0 & 12 & 0 & -30 & 0 \\ 0 & 0 & 0 & 20 & 0 & -70 \\ 0 & 0 & 0 & 0 & 35 & 0 \\ 0 & 0 & 0 & 0 & 0 & 63 \end{bmatrix} \end{array} . \quad (4.10)$$

Two important points are apparent: the transformation is exact in this integer form and the inverse transformation is inexpensive because T_d is an upper triangular matrix.

4.3 Probe functions

The first example extols the main point; the presumption of orthogonality destroys performance and provides an inferior result. Next, a discontinuous function reveals the computational troubles encountered outside of Riesz–Fischer.

4.3.1 Example I: smooth function

Start with the amplitude vector $\alpha^T = (0, 1, 2, 3, 0, 0, 0)$ to define a C^1 function,[3]

$$\begin{aligned} f(x) &= \alpha \cdot \{P_n(x)\}_{n=0}^6 \\ &= P_1(x) + 2P_2(x) + 3P_3(x) \\ &= \frac{1}{2}\left(-1 - x + 3x^2 + 5x^3\right). \end{aligned} \quad (4.11)$$

Problem– What problems arise from the assumption that the Legendre polynomials are orthogonal over $l^2[-1, 1]$? Equation (4.6) is used in a corrupted form,

$$a_k = \frac{2k+1}{2}\Delta \sum_{x \in \sigma} f(x)P_k(x), \quad (4.12)$$

where Δ represents a uniform mesh spacing. The results are shown in Figure 1. The solid symbols are for amplitudes of input terms which are nonzero, open symbols indicate spurious modes created by using an inappropriate numerical method. The method converges linearly, which implies that reasonable accuracy is very expensive. Spurious modes are created, modes with amplitude 0 have nonzero values.

Solution– Switch to the monomial basis $\left(g_k(x) = x^k\right)$, and compute the amplitudes using (3.10). Use an affine transformation to recover the Legendre amplitudes, for example $b = T_6 a$. The algorithm is exact. The test case used $\Delta = 0.25$ for a double precision computation in *Mathematica* and recovered exact answers[4] for $\alpha_0 - \alpha_6$.

Source of error– The root of the error is analytic integrals do not equal their numeric counterparts; the off-diagonal elements in (3.10) are not zero and the solution methods

---

[3]Functions with a continuous first derivative.

[4]The exact answer is attributed to the twin blessings of a high quality linear solver and the x values having exact binary representation.
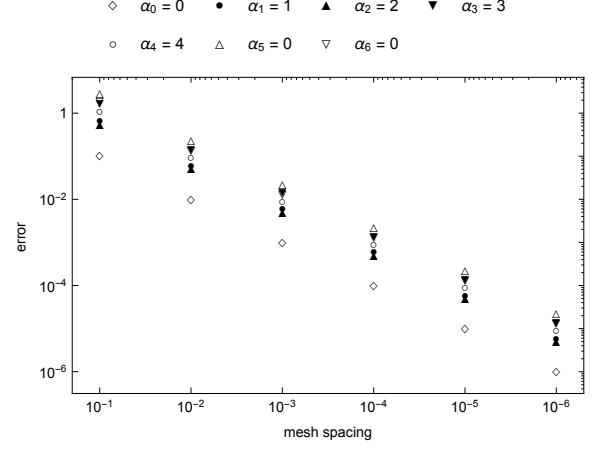

Fig. 1: Error in computing the Legendre amplitudes for the function in (4.11).

for (3.12) don't work. For example, the product function $P_0(x)P_2(x)$, which corresponds to $\mathbf{A}^*\mathbf{A}$ matrix element $(r, c) = (1, 3)$ in (3.3),

$$\int_{-1}^1 P_0(x)P_2(x)\,dx \neq \sum_{\nu=0}^{2n} P_0(x_\nu)P_2(x_\nu)\Delta. \quad (4.13)$$

The left-hand side is an analytic integral exactly equal to 0 and the right-hand side is a numerical summation. For $\Delta = \frac{1}{4}$ the integral is approximated as

$$\sum_{\nu=1}^{8}\frac{1}{2}\left(3x_\nu^2 - 1\right)\Delta = \frac{1}{64} \neq 0,$$

a statement of orthogonality lost. Figure 2 contrasts the two integration methods.

4.3.2 Example II: discontinuous function

Stepping outside of the strictures of the Riesz–Fischer theorem, consider a discontinuous function,

$$\operatorname{sgn} x = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}. \quad (4.14)$$

The sign function has odd parity so the approximation contains only odd terms.

Figure 3 compares fits with maximum order $d = 50$, $d = 100$, and $d = 200$ against the sign function. The amplitudes for the approximations are shown in Figures 4 and 5.

Amplitudes for the monomial approximations are shown in Figure 4. To use a logarithmic scale, the absolute value of the amplitude $|a|$, is shown. Filled rectangles indicate the sign of the amplitude is positive, open rectangles indicate negative values. Amplitudes for even functions have value 0 and are not plotted. The amplitudes grow without bound.
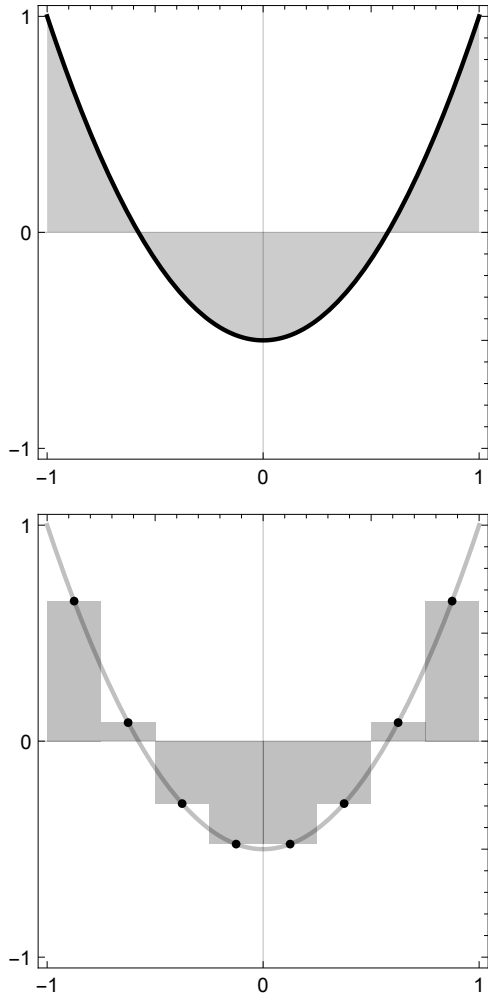
Fig. 2: Error in computing matrix cross terms. The function is $P_0(x)P_2(x)$.

Notice the invariance in form of the amplitudes; the scale is changing, but not the shape.

Amplitudes for the Legendre approximations are shown in Figure 5. Positive values are shown with filled rectangles, 0 values with open diamonds, and negative values with open rectangles. These amplitudes are bounded. The plots demonstrate that the values of the lowest order terms do not change as the order of fit is increased. Notice the invariance of amplitude values; the lower order amplitudes do not change as the fit order increases.

Problem– Monomial amplitudes grow without bound quickly exhausting the range of double precision computation. In such cases, orthogonal polynomials provide infinity insurance; Legendre amplitudes are bound. The monomial linear system is ill-conditioned. After all, these functions are merely linearly independent. But the Legendre polynomials present a much better conditioned linear system as they are closer to being orthogonal in l^2.

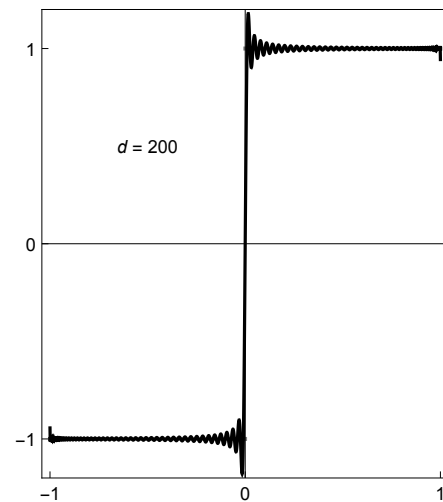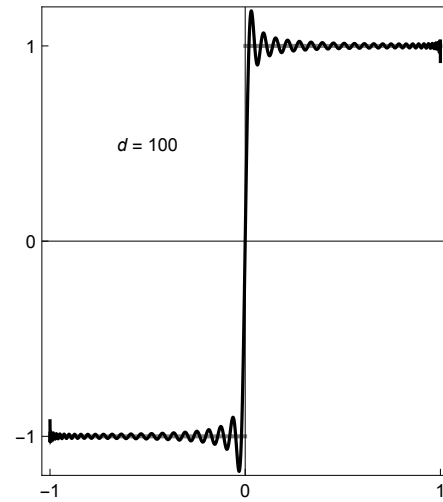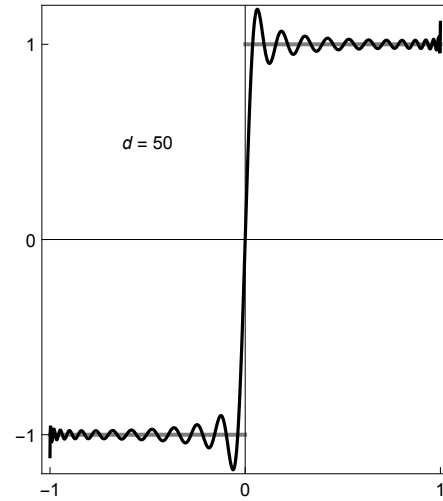**Solution–** Use the Legendre basis $(g_k(x) = P_k(x))$ and

Fig. 3: Approximation of the sign function showing the Gibbs phenomena at the midpoint and boundaries.

Fig. 4: Amplitudes for the monomial approximation of the sign function in (4.14) found by solving (3.10).
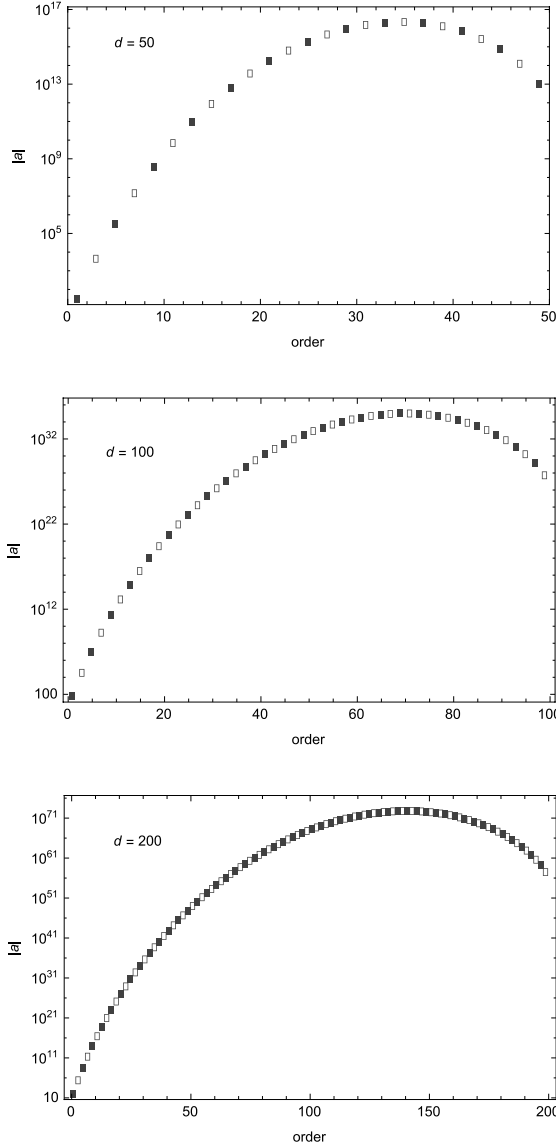

Fig. 5: Amplitudes for the Legendre approximation of the sign function in (4.14) found by solving (3.10).

compute the amplitudes using (3.10). Affine transformations recover the monomial amplitudes, for example $a = T_{200}^{-1}b$.

5. Conclusion

Computation demands discrete evaluation and finite meshes. Outside of sine and cosine, there are no functions continuous in both L^2 and l^2. The presumption of orthogonality in discrete space introduces extreme numerical errors. Instead, rely upon linear independence to create and solve the linear systems. Affine transformations connect the function spaces of choice. When resolving discontinuous functions, L^2-orthogonal functions can create a dense linear system with improved conditioning.
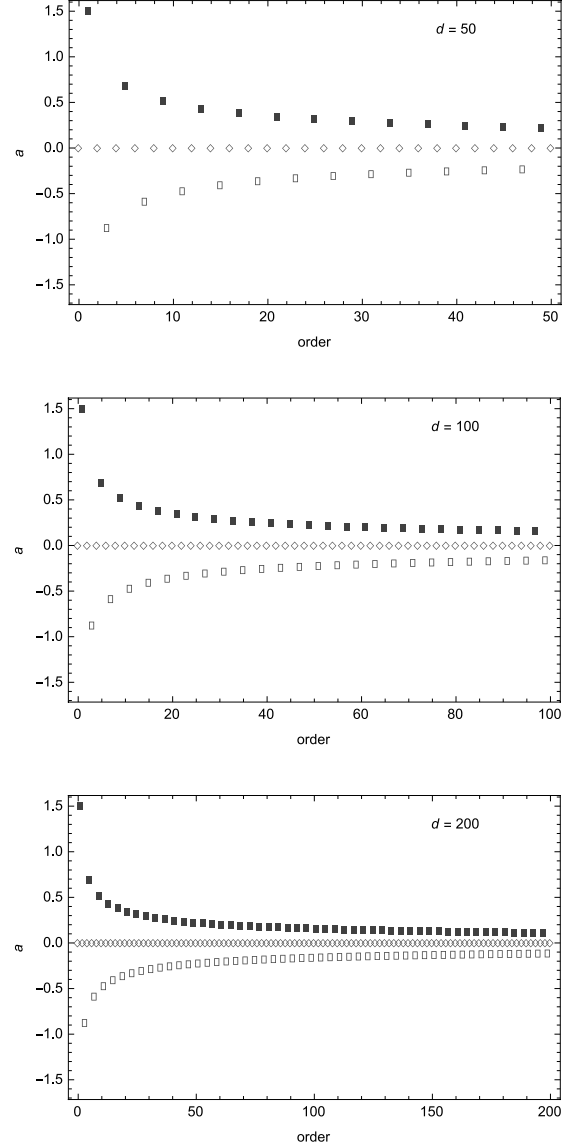
References

[1] Bevington, P. R., *Data Reduction and Error Analysis for the Physical Sciences* 1e, McGraw Hill, 1969.
[2] Rudin, W., *Principles of Mathematical Analysis* 3e, McGraw-Hill, 1976.
[3] Meyer, C. D., *Matrix Analysis and Applied Linear Algebra*, SIAM, 2004.
[4] Laub, A. J., *Matrix Analysis for Scientists and Engineers*, SIAM, 2005.
[5] Weisstein, E.W., "Legendre Polynomial." MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LegendrePolynomial.html