Exponential grids in high-dimensional space

Peter R. Brune
Department of Computer Science, University of Chicago
Matthew G. Knepley
The Computation Institute, University of Chicago
L. Ridgway Scott
Departments of Computer Science and Mathematics, University of Chicago
Chicago, Illinois, USA

December 15, 2011

Abstract

We consider the approximation of functions that are localized in space. We show that it is possible to define meshes to approximate such functions with the property that the number of vertices grows only linearly in dimension. In one dimension, we discuss the optimal mesh for approximating exponentially decreasing functions. We review the use of Cartesian product grids in multiple dimensions introduced in a paper of Bank and Scott [4].

Keywords: exponential grids, Cartesian product grids, geometric grids, curse of dimensionality

1 Introduction

Many approximation problems suffer the 'curse of dimensionality' such as occurs in quantum mechanical models of atomic systems [15]. On the other hand, in these systems the functions being approximated decay exponentially away from a small number of atomic centers. This leads to the possibility of approximating them in configuration space by exponentially (or geometrically) graded meshes that could potentially reduce the number of degrees of freedom drastically. We present here a simple model problem to illustrate the ideas.

2 Example: one dimension

The basis for all the meshes studied here is a simple one-dimensional grid for an interval $[0, X]$:

$$0 = x_0 < x_1 < \cdots < x_N = X. \tag{2.1}$$

We will consider graded meshes, in which the mesh size $x_{i+1} - x_i$ increases as i increases.

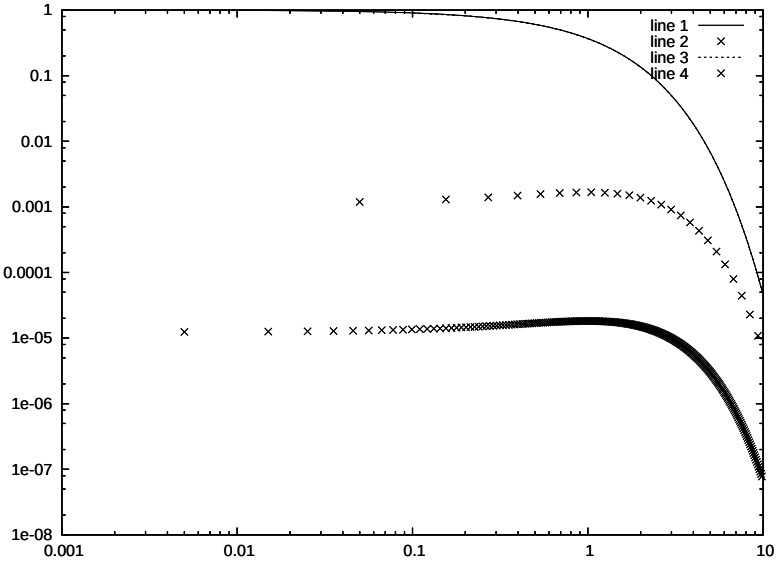

Figure 1: Approximation of e^{-r} by piecewise linears on a one-dimensional geometrically graded mesh. The horizontal axis is the variable r. The top curve is e^{-r} and the bottom 'curves' represent the errors for $h = \delta = 0.1$ (upper, 25 points) and $h = \delta = 0.01$ (lower, 240 points), where $\gamma = 1 + \delta$ in both cases. Note the concentration of interpolation points near the 'knee' of the exponential curve.

2.1 Geometric meshes

One familiar mesh of this type is the geometric mesh:

$$x_{i+1} - x_i = h\gamma^i,\ i = 0, \ldots, N, \tag{2.2}$$

where $x_0 = 0$ and $\gamma > 1$. Thus we see that

$$x_k = \sum_{i=0}^{k-1}(x_{i+1} - x_i) = h\sum_{i=0}^{k-1}\gamma^i = h\frac{\gamma^k - 1}{\gamma - 1}. \tag{2.3}$$

In particular, $x_N = h\frac{\gamma^N - 1}{\gamma - 1} = X$ provided that

$$N = \frac{\log\left(1 + (X/h)(\gamma - 1)\right)}{\log\gamma}. \tag{2.4}$$

Also, note that the mesh increments are related to the distance from the origin by

$$h_k := x_{k+1} - x_k = (\gamma - 1)x_k + h = (\gamma - 1)x_k + x_1, \tag{2.5}$$

where $h = h_0$. Also, if $\gamma = 1 + ch$, then the number of mesh points N in (2.4) is given by

$$N = \frac{\log\left(1 + cX\right)}{\log(1 + ch)} \approx \frac{\log\left(1 + cX\right)}{ch}. \tag{2.6}$$

Now suppose we consider approximating $u(x) = e^{-x}$ by piecewise linear functions on a geometric mesh. The choice $h = \gamma - 1$ is close to optimal in the approximation depicted in

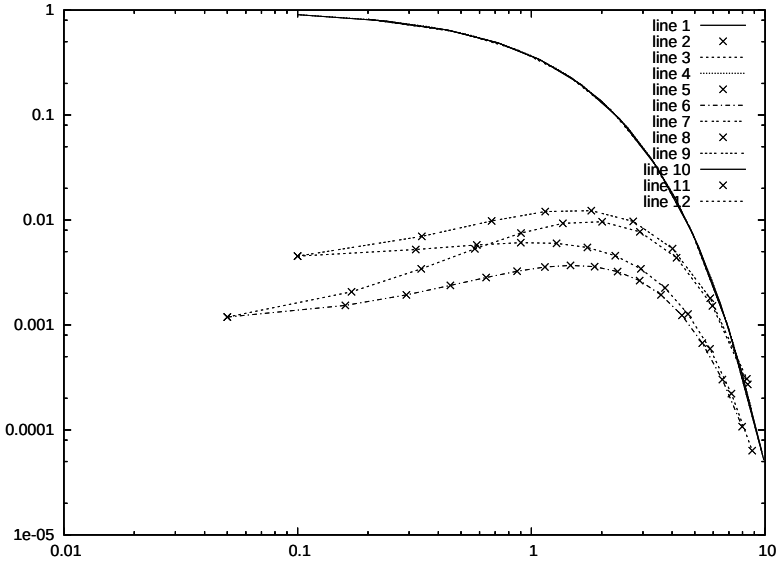

Figure 2: Approximation of e^{-r} by piecewise linears on a series of one-dimensional geometrically graded meshes. The horizontal axis is the variable r. The top curve is e^{-r} and the bottom 'curves' represent the errors for $(h, \gamma) = (.2, 1.2), (.1, 1.2), (.1, 1.4), (.2, 1.4)$. Note the error increase near the 'knee' of the exponential curve.

Figure 1, cf. Figure 2. However, we also see that the error (shown in the lower curves for two different choices of h) are not optimal, in that the errors are not constant. There are points toward the end of the interval that do not contribute materially to the approximation, and there is a bulge in the error in the part of the interval just to the right of $x = 1$. This bulge is exaggerated if one chooses $\gamma \neq 1 + h$, as is indicated in Figure 2.

2.2 Exponential meshes

We now ask what the form of the optimal mesh for approximating a decaying exponential might be. We begin with the piecewise linear case. As before, we expect the error on each subinterval to be of order h_i^2, where $h_i = x_{i+1} - x_i$, $i = 1, \ldots, n$. The interpolation error can be approximated by the error at the midpoint of each interval, which can be estimated by Taylor's theorem as

$$\begin{aligned} & u((x_i + x_{i+1})/2) - (u(x_i) + u(x_{i+1}))/2 \\ =& \frac{1}{8} (x_{i+1} - x_i)^2 u''((x_i + x_{i+1})/2) + \mathrm{O}\left((x_{i+1} - x_i)^4\right). \end{aligned} \tag{2.7}$$

Suppose that we want to equilibrate the error at a level h^2 on each interval for the function $u(x) = e^{-x}$. Here we are thinking of h as a resolution parameter. For a uniformly complicated function, we expect an error of the form Ch^2 for a uniform mesh of size h. Now we use this as a guide to get a uniform error for a function whose complexity varies exponentially.

Dropping high-order terms, we thus want

$$\begin{aligned} h^2 =& \frac{1}{8} \left(x_{i+1} - x_i\right)^2 u''((x_i + x_{i+1})/2) \\ =& \frac{1}{8} \left(x_{i+1} - x_i\right)^2 e^{-((x_i+x_{i+1})/2)}. \end{aligned} \tag{2.8}$$

Taking square roots we find

$$h = \frac{\sqrt{2}}{2} \left(x_{i+1} - x_i\right) e^{-((x_i+x_{i+1})/4)}. \tag{2.9}$$

Setting $x_0 = 0$, we can view this as a difference method (trapezoidal rule)

$$\frac{(x_{i+1} - x_i)}{h} = \sqrt{2} e^{((x_i+x_{i+1})/4)} \tag{2.10}$$

for the differential equation

$$x' = \sqrt{2} e^{x/2} \tag{2.11}$$

whose solution blows up in finite time. In particular,

$$e^{x(t)/2} = \left(e^{-x(0)/2} - t/\sqrt{2}\right)^{-1}. \tag{2.12}$$

This implies that, for any given resolution h, only a finite domain need be considered. One series of such meshes is depicted in Figure 3.

Unfortunately, the number of grid points in an exponential mesh grows inversely to the resolution parameter h. In the error curves in Figure 3, the number of mesh points is doubling (from 7 to 14 to 28) as the resolution is decreased. This behavior can be verified for higher resolution as well:

$$h \approx 1.4/N. \tag{2.13}$$

The error is quadratic, of the order h^2, but still we require a substantial mesh size (of order $1/h$) to resolve e^{-r} accurately with a fixed degree of approximation. The main concentration of meshpoints is again near the knee in the curve e^{-r}.

2.3 Higher-order approximations

Suppose we now allow variable order approximations. The results for linear approximations both on geometric meshes and on optimal exponential mesh suggest that greater accuracy would be beneficial in the knee of the exponential curve.

To estimate the error in approximating e^{-r} via polynomials of arbitrary degree, we simply use Taylor's theorem. Thus the error on the i-th mesh interval $[x_i, x_{i+1}]$ would be approximately

$$\epsilon_i = \frac{h_i^{k_i}}{k_i!} e^{-x_i}, \tag{2.14}$$

for polynomials of degree $k_i - 1$ where $h_i = x_{i+1} - x_i$.

To get a feeling for how variable approximation order affects the computations, we consider a family of geometric meshes. The top curves in Figure 4 depict the order k of approximation required to keep the error below a fixed tolerance; $k = 2$ corresponds to linear

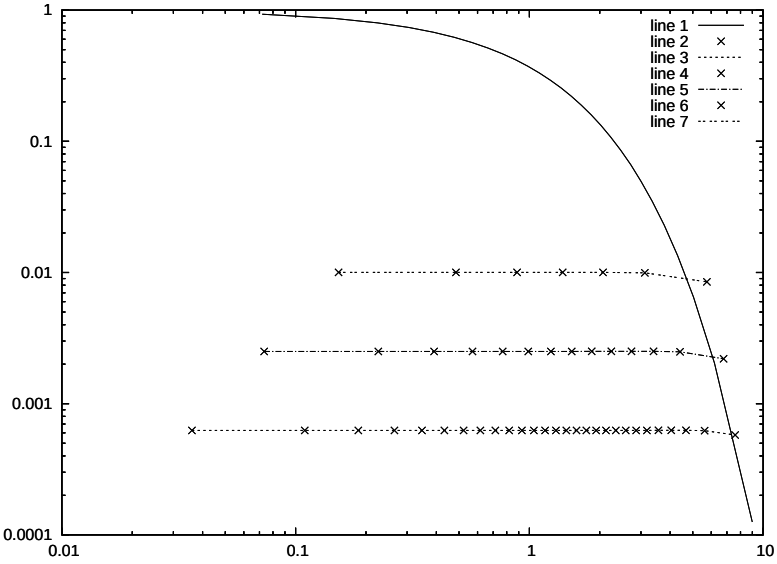

Figure 3: Approximation of e^{-r} by piecewise linears on a series of one-dimensional exponential meshes. The horizontal axis is the variable r. The top curve is e^{-r} and the bottom 'curves' represent the errors for three exponential meshes.

approximation. The middle curves just depict e^{-r} for convenience, whereas the bottom curves are the error estimates in (2.14). We see that very large approximation order k can be required in some cases. Moreover, the order can be low initially, only growing near the knee in e^{-r}. Surprisingly, the required order continues to grow past the knee, and then it drops to the lowest order (corresponding to piecewise constant approximation). This occurs when the exact curve e^{-r} drops below the error tolerance, so that an approximation by zero would be sufficient.

The geometric mesh leads to overly large elements, requiring higher degrees of approximation than may be required from an efficiency point of view. The optimal choice of mesh size and polynomial degree is known to be difficult. We recall the form of the optimization for completeness.

There are two possible forms of the mesh-degree optimization problem. In the first, we ask what the best mesh and approximation degree assignment is for a given number of mesh points. We can express this as

$$\min_{h_i>0,k_i\geq 0} \max_{i=1,\ldots,n} \epsilon_i = \min_{h_i\geq 0,k_i\geq 0} \max_{i=1,\ldots,n} \frac{h_i^{k_i}}{k_i!} e^{-\sum_{j=1}^{i} h_i}. \tag{2.15}$$

Here the optimization variables $h_1, \ldots, h_n$ range over positive real numbers, and the optimization variables $k_1, \ldots, k_n$ range over nonnegative integers. (One simple relaxation of (2.15) is to allow $k_1, \ldots, k_n$ to range over nonnegative real numbers.) This form of the optimization problem asks for the optimal order of approximation given n mesh intervals, but it does not take into account the fact that the higher degrees involve more work.

A more useful optimization problem involves an estimate $\omega(k_1, \ldots, k_n)$ required to obtain the solution using polynomials of degree k on a given interval. The optimization problem to

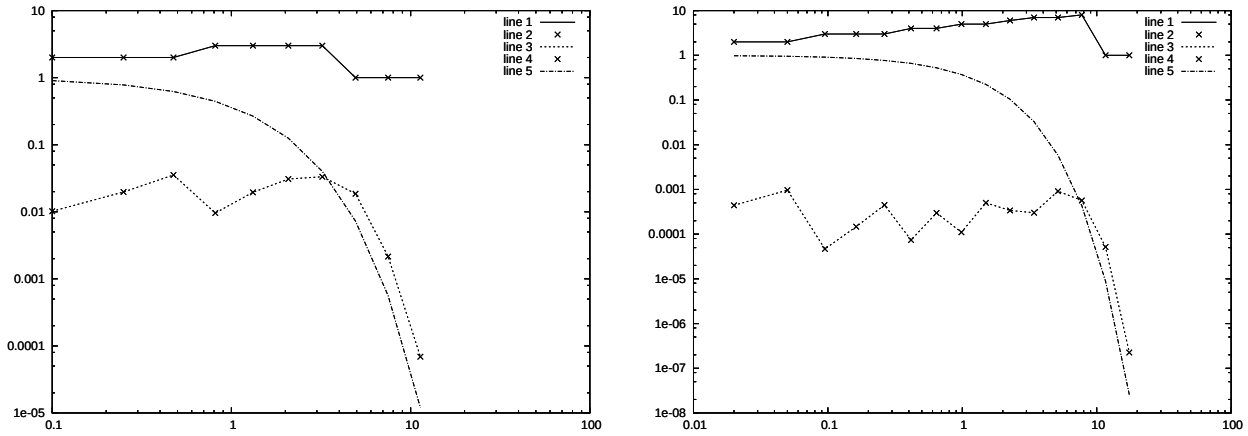

Figure 4: Two examples of the use of variable order approximation on geometric meshes with different tolerances and different initial mesh sizes, based on the error model (2.14). The horizontal axis is the variable r. The top curves give the order k of approximation ($k = 2$ corresponds to linear approximation), the middle curves recall e^{-r}, and the bottom curves are the error estimates in (2.14).

minimize the work to achieve an accuracy $\alpha > 0$ is then

$$\min_{h_i>0, k_i\geq 0} \omega(k_1, \ldots, k_n) \quad \text{subject to} \quad \max_{i=1,\ldots,n} \frac{h_i^{k_i}}{k_i!} e^{-\sum_{j=1}^{i} h_i} \leq \alpha. \tag{2.16}$$

The constraints in (2.16) are clearly more complex than in (2.15).

3 Cartesian product grids

We now generalize the one dimensional meshes by a technique which provides good approximation with limited mesh points at least in low-dimensional cases. It is easy to describe, at least in two dimensions. Given a one dimensional mesh as described in section 2, we can define a mesh in two dimensions via the Cartesian mesh points

$$(0, \pm x_i), (\pm x_i, 0), (\pm x_i, \pm x_i),\ i = 1, \ldots, N. \tag{3.17}$$

One eighth of such a mesh is depicted in Figure 5. Note that there are two basic triangle similarity classes. If the one-dimensional mesh is geometric with $\gamma = \sqrt{2}$, then all of the triangles are similar; this type of mesh was used in [4, 8]. The type of mesh in Figure 5 has also been referred to as a "geometric mesh" [2].

Cartesian product meshes can be defined in any number of dimensions, but the number of mesh points grows exponentially in the number of dimensions. Thus we are led to look for a mesh that has a smaller rate of growth with respect to the dimension.

4 Prismatic elements

Prism elements have been used in various contexts [3, 10, 14, 17]. We will restrict to ones with a simple element domain, and we only consider only the lowest-order case in full detail.

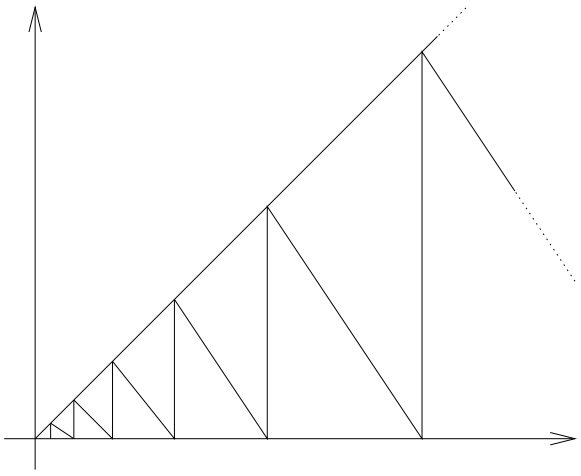

Figure 5: Triangulation of the domain $\widetilde{D} = \{(r_1, r_2) \in \mathbb{R}^2_+ : r_2 \leq r_1\}$ via a Cartesian product grid.

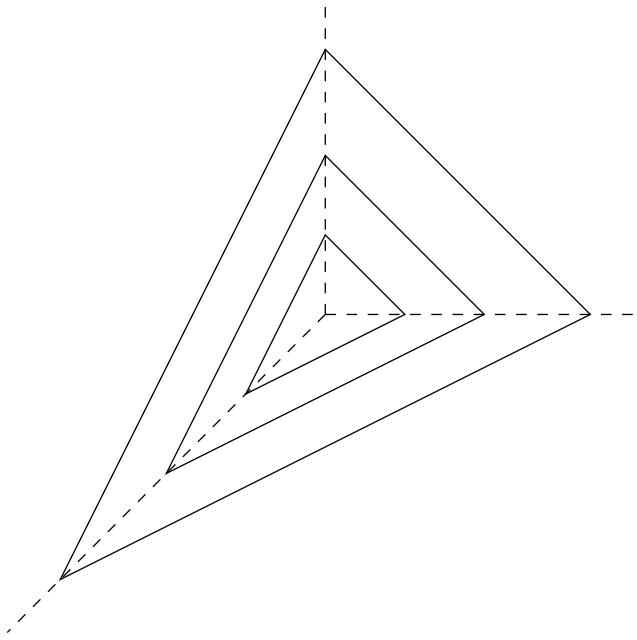

Figure 6: Generation of prismatic meshes in two dimensions; nested similar triangles.

4.1 Prismatic meshes

We now address the problem of finding meshes that subdivide $\mathbb{R}^d$ with a number of mesh points that grows at most linearly in d, and like the mesh (3.17), grows at most logarithmically in the ratio of the domain size to the smallest mesh size.

Let us describe the construction in two dimensions. We begin with two similar triangles centered at the origin, as depicted in Figure 6. We start with the triangles with vertices

$$\begin{aligned}&[(x_1,0),(0,x_1),(-x_1,-x_1)],\ [(x_2,0),(0,x_2),(-x_2,-x_2)],\\&\ldots,[(x_N,0),(0,x_N),(-x_d,-x_N)].\end{aligned} \tag{4.18}$$

The dashed lines connect the vertices with the triangle barycenters. The prisms (quadrilaterals) bounded by these lines and the edges of the triangles form the prismatic elements of the triangulations.

In the general case, we start with a simplex Σ_0 whose barycenter is at the origin. Define

$$\Sigma_i = \gamma_i \Sigma_0 = \left\{ \gamma_i x \;\middle|\; x \in \Sigma_0 \right\}, \tag{4.19}$$

where the scaling factors γ_i satisfy $1 < \gamma_1 < \gamma_2 < \cdots < \gamma_N$. We can write the mesh in terms of annuli of the form $\Sigma_{i+1} \backslash \Sigma_i$. We now describe how these annuli are divided into prismatic domains.

Let $\sigma^1, \ldots, \sigma^{d+1}$ denote the barycentric subdivision of Σ_0. Note that each σ^i can be associated with one face F_i of Σ_0; σ^i is the convex hull of the origin and F_i. Then we can write the prismatic element domains in the mesh as

$$\sigma_{ij} = \gamma_i \sigma^j \backslash \gamma_{i-1} \sigma^j,\ j = 1, \ldots, d+1,\ i = 1, \ldots, N, \tag{4.20}$$

where we add $\gamma_0 = 1$ and analogous to (4.19), we define

$$\gamma \sigma = \left\{ \gamma x \;\middle|\; x \in \sigma \right\}. \tag{4.21}$$

4.2 Element domains

We begin by defining the reference domain(s). Let Σ denote the reference simplex

$$\Sigma = \left\{ \xi \in \mathbb{R}^d \;\middle|\; 0 \le \xi_i,\ i = 1, \ldots, d \quad \text{and} \quad \xi_1 + \cdots + \xi_d \le 1 \right\}. \tag{4.22}$$

We consider prisms that can be written as a truncated simplex:

$$\Sigma_H = \left\{ \xi \in \Sigma \;\middle|\; \xi_d \le H \right\}, \tag{4.23}$$

where H is a parameter in the interval $0 < H < 1$, which we call the *aspect ratio* of the prism.

Note that for each prism reference domain Σ_H there is a distinguished point $p = (0, \ldots, 0, 1)$ (outside the domain) that represents the point of intersection of the edges of Σ_H which have the ξ_d coordinate varying.

The set of element domains we will consider consist of affine maps of a fixed Σ_H for a fixed H. A typical element domain is depicted in Figure 7 in the three dimensional case. For each such element domain $e = J\Sigma_H + \hat{\xi}$, where J is the Jacobian of the affine map, there is a distinguished point $p_e = Jp + \hat{\xi} = J_d + \hat{\xi}$ where J_d is the d-th column of the matrix J. Note also that the faces of the prisms having such edges are themselves either simplices of dimension $n-1$ or prisms of dimension $n-1$ with the same aspect ratio H. All of the prismatic faces of a given element e have the same distinguished point p_e as the parent prism.

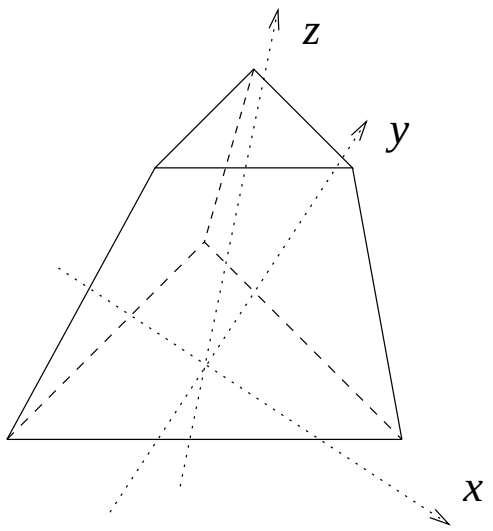

Figure 7: A three-dimensional prism. The smaller triangle is the top, and the larger triangle is the base.The dashed lines are at the back of the prism.

4.3 Element functions

In the simplest case, prism elements [10, 14] have a natural set of basis functions which arise as a tensor product of linear polynomials in $d-1$-dimensional variables and linear functions in one variable. There is a base and top to each pyramid. This is depicted in Figure 7 in the three dimensional case. Let us assume that the base lies in the plane $(x_1, \ldots, x_{d-1}, 0)$ and the top lies in the plane $(x_1, \ldots, x_{d-1}, H)$. Then the functions are

$$\Pi_k^d = \left\{ (a + bx_d) p(x_1, \ldots, x_{d-1}) \;\middle|\; a, b \in \mathbb{R},\ p \in \mathcal{P}_k[x_1, \ldots, x_{d-1}] \right\}. \tag{4.24}$$

For $d = 2$, Π_1^d consists of bilinear functions $Q_1(x_1, x_2)$, and for any d we have

$$\mathcal{P}_k[x_1, \ldots, x_d] \subset \Pi_k^d \subset Q_k[x_1, \ldots, x_d]. \tag{4.25}$$

But the dimension of $\Pi_1^d = 2 \dim \mathcal{P}_1[x_1, \ldots, x_{d-1}] = 2d$, whereas the tensor-product space of d-linear functions $Q_1[x_1, \ldots, x_d]$ has dimension 2^d.

4.4 Continuity of elements

Note that the restrictions of the element functions to the prismatic faces of each element domain are themselves element functions of one lower dimension. Therefore they form continuous elements on a prismatic mesh as described in Section 4.1.

4.5 Number of vertices and prisms

Suppose that N is the number of the simplices involved in the mesh. For example, Figure 6 depicts the case $N = 3$ in $d = 2$ dimensions. In general, we have N similar d-dimensional simplices with the same barycenter that are nested in a fashion similar to Figure 6, each one bigger than the last. We may think of these simplices being a mathematical analog of the familiar nested Russian dolls. Let us number the simplices $\Sigma_0, \ldots, \Sigma_N$ where Σ_0 is the

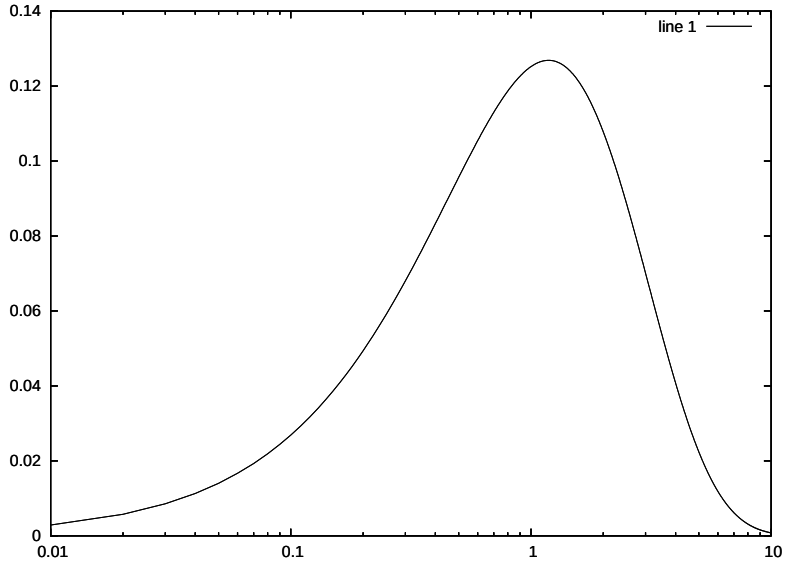

Figure 8: Graph of the function ψ defined in (5.27).

smallest and Σ_N is the largest. In each annulus $A_j = \Sigma_{j+1} - \Sigma_j$, $j = 1, \ldots, N-1$, there are $d+1$ prisms, one for each face of Σ_{j+1} and Σ_j that forms the top and the base, respectively, of a prism (there are $d+1$ faces of a d-simplex).

The set of vertices in the prismatic mesh is the same as the number of vertices in the union of the simplices $\Sigma_1, \ldots, \Sigma_N$. The number of vertices of a d-simplex is $d+1$. Thus the number of vertices in the prismatic mesh is $N(d+1)$.

The number of prisms that arise in the subdivision of each annulus increases with the dimension. For each face of a d-dimensional simplex (of which there are $d+1$), there is an associated prism in each annulus. Thus there are $(d+1)(N-1)$ prisms. In the assembly process, each prism contributes to $2d$ vertices on the top and bottom of each prism. Thus the standard assembly algorithm will take $\mathrm{O}(d^2)$ work.

5 Approximation of exponentially decreasing functions

Unfortunately, approximation of exponentially decreasing functions in multiple dimensions is more complex than in one dimension. As an example, consider a trapezoidal mesh in n dimensions with mesh points $(t_k, 0, \ldots, 0)$, $(0, t_k, 0, \ldots, 0)$, $\ldots$, $(0, \ldots, 0, t_k)$. We can write these points more succinctly as

$$t_k \mathbf{e}^i, \ i = 1, \ldots, n, \tag{5.26}$$

where $\mathbf{e}^i$ is the canonical basis vector: $\mathbf{e}^i_j = \delta_{ij}$.

The interpolant based on these mesh points is constant along any edge connecting these points, so the error, say, at $\mathbf{x}_k = \frac{1}{2}(t_k, t_k, 0, \ldots, 0)$ satisfies

$$(u - u_I)(\mathbf{x}_k) = e^{-t_k/\sqrt{2}} - e^{-t_k} = \psi(t_k) \tag{5.27}$$

for the case when u_I denotes linear interpolation. When t is small, $\psi(t)$ is small, and when t is large, $\psi(t)$ is again small. But in between, $\psi(t)$ is not small, as depicted in Figure 8.

This implies that additional angular approximation is needed in the intermediate region, similar to the 'knee' area of the exponential where approximation is known to be criticial from Figures 2 and 3.

It is useful to see what standard error estimates for linear approximation would yield to confirm the failure of this simple approach. It is possible to derive standard error estimates suitable for exponentially decreasing functions on geometric grids. On any element e, we have the following estimate [18] for $1 \leq q \leq \infty$ and $m = 0, 1$:

$$\| u - u_I \|_{W_q^m(e)} \leq C \mathrm{diam}(e)^{2-m} \| u \|_{W_q^2(e)}. \tag{5.28}$$

For simplicity, let us assume that the radial decomposition is based on a geometric subdivision (2.2), namely, $t_{k+1} - t_k = \gamma^k h$. Thus

$$t_k = h \frac{\gamma^k - 1}{\gamma - 1}. \tag{5.29}$$

Suppose that e lies in the k-th ring from the origin. Then

$$\mathrm{diam}(e) \leq C h \gamma^k. \tag{5.30}$$

The points in the prismatic face of e closest to the origin are a convex combination of vertices $v_1, \ldots, v_d$ of that face. Recall that $v_i = t_k e^i$ for suitable numberings of the dimensions. The closest point on this face to the origin is

$$\frac{1}{d} \sum_{i=1}^{d} v_i = \frac{t_k}{d} (1, 1, \ldots, 1, 0) \tag{5.31}$$

in these coordinates. The distance from this point to the origin is $t_k / \sqrt{d}$. Thus

$$\min \left\{ |x| \;\middle|\; x \in e \right\} \geq \frac{t_k}{\sqrt{d}} = \frac{h}{\sqrt{d}} \frac{\gamma^k - 1}{\gamma - 1}. \tag{5.32}$$

Define $\nabla_2 u(x)$ to be the tensor of all second derivatives of u, and we let $|\nabla_2 u(x)|$ denote some suitable norm on this tensor. Suppose that

$$|\nabla_2 u(x)| \leq C e^{-\alpha |x|} U(x) \tag{5.33}$$

for all $x \in \mathbb{R}^d$, where α is a positive real number that encodes the rate of decrease of u, and U just accounts for the difference between $|\nabla_2 u(x)|$ and a pure exponential decrease. Then standard interpolation estimates yield (for either $m = 0$ or $m = 1$)

$$\begin{aligned} \| u - u_I \|_{W_q^m(e)} \leq & C (h \gamma^k)^{2-m} \| u \|_{W_q^2(e)} \qquad \text{[by (5.30)]} \\ \leq & C (h \gamma^k)^{2-m} e^{-\frac{\alpha h}{\gamma - 1} (\gamma^k - 1)} \| U \|_{L^q(e)} \qquad \text{[by (5.32) and (5.33)]} \\ = & C h^{2-m} e^{(2-m) \log(\gamma^k) - \frac{\alpha h}{\gamma - 1} (\gamma^k - 1)} \| U \|_{L_q(e)}, \end{aligned} \tag{5.34}$$

where here and subsequently constants C may depend on the dimension d. We see that the parameter

$$\rho = \frac{\alpha h}{\gamma - 1} \tag{5.35}$$

plays a key role in the estimate (5.34). We can turn this around to say that

$$\gamma = 1 + (\alpha/\rho)h. \tag{5.36}$$

The expression $(2-m)\log t - \rho(t-1)$ has its maximum at $t = (2-m)/\rho$, so that

$$e^{(2-m)\log t - \rho(t-1)} \leq \left(\frac{2-m}{\rho}\right)^{2-m} e^{-(2-m)+\rho}. \tag{5.37}$$

Therefore

$$\begin{aligned} \| u - u_I \|_{W_q^m(e)} &\leq Ch^{2-m} \left(\frac{2-m}{\rho}\right)^{2-m} e^{m-2+\rho} \| U \|_{L_q(e)} \\ &= C \left(\frac{(\gamma-1)(2-m)}{\alpha}\right)^{2-m} e^{m-2+\rho} \| U \|_{L_q(e)}. \end{aligned} \tag{5.38}$$

The expression

$$\psi(\rho) = \left(\frac{2-m}{\rho}\right)^{2-m} e^{-(2-m)+\rho} \tag{5.39}$$

has its minimum $\psi_{\min} = 1$ at $\rho = 2-m$, and $\psi(2-m) = 1$. This suggests that the optimal choice is to take $\rho = 2-m$, and thus by (5.36) we have

$$\gamma = 1 + (\alpha/(2-m))h, \tag{5.40}$$

and thus

$$h = (2-m)(\gamma-1)/\alpha. \tag{5.41}$$

In this case, choosing the optimal γ, (5.38) simplifies to

$$\| u - u_I \|_{W_q^m(e)} \leq Ch^{2-m} \| U \|_{L_q(e)}. \tag{5.42}$$

Unfortunately, (5.40) implies that the total number of mesh points will be proportional to h^{-1}, cf. (2.4).

The error expression (5.27) suggests that the use of exponential grids in multiple dimensions requires the use of higher-degree elements to resolve the angular dependence in at least some part of the domain. If we use degree k elements, then the number of degrees of freedom is no longer linear in dimension. Higher-order elements have degrees of freedom on edges, faces and so forth. The number of edges increases quadratically in the dimension for an exponential mesh, and the number of faces increases cubically. Thus the number of degrees of freedom for k-th order elements will be of order k in the dimension for an exponential mesh. However, if we take an $h-P$ strategy [2], and let the degree k be large only where needed, we still have the potential to have an effective algorithm. We indicate the feasibility of such an approach by example.

6 Example: hydrogen atom

The hydrogen atom [15] can be modeled as a solution to (6.46), cast in terms of an eigenproblem

$$\Delta_3 u(\mathbf{r}) + \frac{2}{|\mathbf{r}|} u(\mathbf{r}) = \lambda u(\mathbf{r}). \tag{6.43}$$

order	radial levels	DoFs
2	40	815
3	40	2435
4	25	3469
5	15	4021

Table 1: Polynomial order, number of radial approximations and total number of degrees of freedom (DoFs) for the hydrogen problem depicted in Figure 9.

One solution is $u(\mathbf{r}) = e^{-|\mathbf{r}|}$ and $\lambda = 1$, which can be seen as follows.

We consider the function of $n = 3$ variables

$$u(\mathbf{r}) = u(r_1, r_2, r_3) = e^{-|\mathbf{r}|} = e^{-\sqrt{r_1^2+r_1^2+r_1^2}}. \tag{6.44}$$

Recall that in spherical coordinates in three dimensions

$$\Delta_3 u(\mathbf{r}) = \frac{\partial^2}{\partial r^2}u + \frac{2}{r}\frac{\partial}{\partial r}u = e^{-|\mathbf{r}|}\left(1 - \frac{2}{|\mathbf{r}|}\right). \tag{6.45}$$

Therefore we can view u alternatively as solution to

$$\Delta_3 u(\mathbf{r}) + \frac{2}{|\mathbf{r}|}u(\mathbf{r}) = f(\mathbf{r}) \tag{6.46}$$

where $f(\mathbf{r}) = e^{-|\mathbf{r}|}$, or to

$$\Delta_3 u(\mathbf{r}) = f(\mathbf{r}) \tag{6.47}$$

where $f(\mathbf{r}) = e^{-|\mathbf{r}|}\left(1 - \frac{2}{|\mathbf{r}|}\right)$. In either case, we must deal with weakly singular integrals in the finite element implementation [13].

If we use a mesh like shown in Figure 6 with only the linear prismatic element, then the best that can happen is convergence to some subspace of functions that is angularly constrained. Indeed, we observed quadratic convergence for the first eigenvalue of (6.43) to a slightly shifted value.

To get full convergence to the eigenproblem, higher-order prismatic elements are required. Figure 9 indicates convergence as the degree is increased. Now we see that convergence can be obtained by increasing the degree of approximation. Examination of the eigenvector approximation shows that the higher-degree approximation is limited to a small part of the domain, suggesting that an $h - P$ approach could be even more efficient.

In full disclosure, we should say that the data in Figure 9 actually represent an overshoot of the first eigenvalue. For a Ritz-Galerkin method, the eigenvalue approximation should be one-sided (from above). However, to evaluate terms involving the singular kernel in (6.43), we used a numerical quadrature scheme that violates the Ritz-Galerkin eigenvalue monotonicity. For this reason, we now discuss a possible approach to the evaluation of the corresponding weakly singular integrals.

A key result depicted in Figure 9 is the resolution of multiple eigenvalues for the excited states of the hydrogen atom. In addition to the convergence of the ground state (lowest eigenvalue, but largest value in Figure 9), the next eigenvalue has a multiplicity of 4.

As the degree of approximation increases, this group of 4 eigenstates gets more clearly resolved in the depiction of the multiple eigenvalues for the approximate system. Similarly, the third eigenvalue has a multiplicity of 9, and these states are well resolved by degree 5. The fourth eigenvalue has a multiplicity of 16, and we can see two groups of eigenvalues (one group of 10, another group of 6), but they have not yet merged at this level of approximation. The grid size is not fixed in Figure 9; rather it becomes more course for higher degree approximation as shown in Table 1.

7 Related work

The notion of 'sparse grids' [5, 11] is closely related to what we study here. Other techniques have also been proposed for problems of high dimension [6, 7, 12].

8 Conclusions

We showed that an $h-P$ strategy could be an effective technique for approximating functions that are localized in space in high dimensions. In one dimension, we derived the optimal mesh for approximating exponentially decreasing functions. We reviewed the use of Cartesian product grids [4] in multiple dimensions. We showed that the curse of dimension can potentially be dispelled (or exorcized) using prismatic elements on appropriate grids.

References

[1] Atkinson, K. Quadrature of singular integrands over surfaces. *Electronic Transactions on Numerical Analysis 17* (2004), 133–150.

[2] Babuska, I., and Guo, B. Q. Approximation properties of the h-p version of the finite element method. *Computer Methods in Applied Mechanics and Engineering 133*, 3-4 (1996), 319 – 346.

[3] Bank, R. E., and Santos, R. F. Analysis of some moving space-time finite element methods. *SIAM Journal on Numerical Analysis 30*, 1 (1993), 1–18.

[4] Bank, R. E., and Scott, L. R. On the conditioning of finite element equations with highly refined meshes. *SIAM J. Num. Anal. 26* (1989), 1383–1394.

[5] Barthelmann, V., Novak, E., and Ritter, K. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics 12* (Mar 2000), 273–288. 10.1023/A:1018977404843.

[6] Beylkin, G., and Mohlenkamp, M. J. Numerical operator calculus in higher dimensions. *Proceedings of the National Academy of Sciences 99*, 16 (2002), 10246–10251.

[7] Beylkin, G., and Mohlenkamp, M. J. Algorithms for numerical analysis in higher dimensions. *SIAM J. Sci. Comput. 26*, 6 (2005), 21332159.

[8] Brenner, S. C., and Scott, L. R. *The Mathematical Theory of Finite Element Methods*, third ed. Springer-Verlag, 2008.

[9] Duffy, M. Quadrature over a pyramid or cube of integrands with a singularity at a vertex. *SIAM journal on Numerical Analysis 19*, 6 (1982), 1260–1262.

[10] Graglia, R., Wilton, D., Peterson, A., and Gheorma, I.-L. Higher order interpolatory vector bases on prism elements. *Antennas and Propagation, IEEE Transactions on 46*, 3 (Mar 1998), 442–450.

[11] Griebel, M., and Oswald, P. On additive Schwarz preconditioners for sparse grid discretizations. *Numerische Mathematik 66* (Dec 1993), 449–463. 10.1007/BF01385707.

[12] Hackbusch, W., Khoromskij, B. N., and Tyrtyshnikov, E. E. Hierarchical Kronecker tensor-product approximations. *Journal of Numerical Mathematics 13*, 2 (20050101), p119 – 156.

[13] Johnson, C., and Scott, L. R. An analysis of quadrature errors in second–kind boundary integral methods. *SIAM J. Num. Anal. 26* (1989), 1356–1382.

[14] Nédélec, J. C. A new family of mixed finite elements in $\mathbb{R}^3$. *Numerische Mathematik 50* (Jan 1986), 57–81. 10.1007/BF01389668.

[15] Pauling, L., and Wilson, E. B. *Introduction to Quantum Mechanics with Applications to Chemistry*. Dover, 1985.

[16] Schwab, C., and Wendland, W. On numerical cubatures of singular surface integrals in boundary element methods. *Numerische Mathematik 62*, 1 (1992), 343–369.

[17] Zhang, S. Numerical integration with Taylor truncations for the quadrilateral and hexahedral finite elements. *Journal of Computational and Applied Mathematics 205* (Aug 2007), 325–342.

[18] Zhang, S., and Scott, L. R. Finite element interpolation of non–smooth functions satisfying boundary conditions. *Math. Comp. 54* (1990), 483–493.

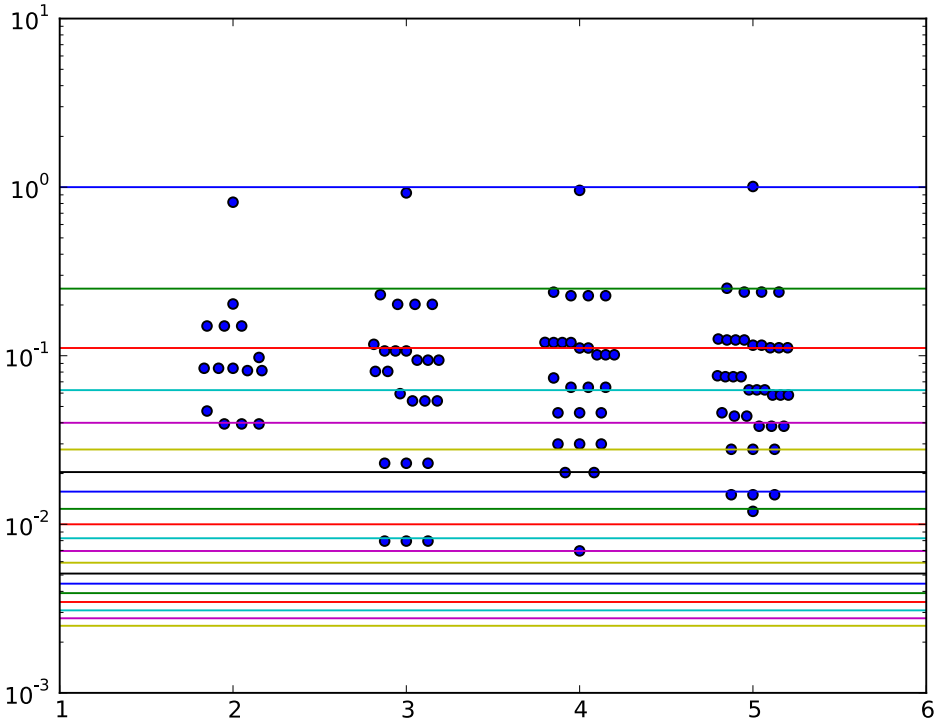

Figure 9: Negative eigenvalues $-\lambda$ for approximations of the Hydrogen atom by direct finite element approximation using prismatic elements for the Schrödinger problem (6.43). The horizontal axis is the degree of polynomial approximation. The mesh descriptors are given in Table 1.