Multiscale Numerical Methods for the Singularly Perturbed Convection-Diffusion Equation

A Thesis by

Peter J. Park

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2000

(Submitted September 8, 1999)

Quod vixi tege, quod vivam rege.

To my parents

Acknowledgements

I am greatly indebted to my advisor Prof. Thomas Hou for his patience and guidance during my time at Caltech. His high standard was an inspiration to me, and he always wanted the best for his students. I am grateful to Dr. Yalchin Efendiev, Dr. Xiaohui Wu, and Prof. Albert Fannjiang for many helpful discussions. They were always generous with their time and provided many insightful comments. I also thank Prof. John Brady, Prof. Joel Franklin, and Dr. Ravi Samtaney for serving on the thesis committee and Dr. Hector Ceniceros for a careful reading of this thesis.

My officemates through the years, Donal Gallagher, Pat Lahey, and Anatoly Baumstein, were much too kind when I bothered them incessantly with questions. Many members of the Applied Mathematics group, including Phill Kang, Maciej Dudek, Russina Philippakos, Jason Kastner, Danny Petrasek, Tom Witelski, Muruhan Rathinam, Postdoctoral Fellows Mike Holst, Anne Gelb, John Pelesko, Patrick Guidotti, Hector Ceniceros, and Professors Don Cohen and Joel Franklin made my stay more pleasant. If it were not for Danny, John, and Patrick, I would not even be addicted to the Red Door coffee. The staff in the department was also wonderful, and I particularly enjoyed my conversations with our sanitary engineer Phil Harter.

Outside the department, I benefitted from numerous dinner conversations with Jerry Shan and Gary Holt regarding the Christian faith and its implications in our lives. I enjoyed my afternoon runs in San Marino with Jerry and learning how to be friends with Linux from Gary. I spent too much time emailing with many friends including Konrad, Ali, Jack, Clara, and Sandra, but I enjoyed every minute of it. Many ophthalmologists tried to help me over the years and I am thankful to them. I must also acknowledge my mentors from the past, in particular my eighth grade algebra teacher Doug Buhler and my undergraduate advisor Howard Stone for their friendship and encouragement. Finally, I thank my sister Jinhee for buying all the clothes I own, and Mom and Dad for never telling me to study.

Abstract

We develop efficient and robust numerical methods in the finite element framework for numerical solutions of the singularly perturbed convection-diffusion equation and of a degenerate elliptic equation. The standard methods for purely elliptic or hyperbolic problems perform poorly when there are sharp boundary and internal layers in the solution caused by the dominant convective effect. We offer a new approach in which we design the finite element basis functions that capture the local behavior correctly.

When the structure of the layers can be determined locally, we apply the multiscale finite element method in which we solve the corresponding homogeneous equation on each element to capture the small scale features of the differential operator. We demonstrate the effectiveness of this method by computing the enhanced diffusivity scaling for a passive scalar in the cellular flow. We carry out the asymptotic error analysis for its convergence rate and perform numerical experiments for verification. When the layer structure is nonlocal, we use a variational principle to gain additional information. For a random velocity field, this variational principle provides correct scaling results. This allows us to design asymptotic basis functions that can capture the global layers correctly.

The same approach is also extended to elliptic problems with high contrast coefficients. When an asymptotic result is available, it is incorporated naturally into the finite element setting developed earlier. When there is a strong singularity due to a discontinuous coefficient, we construct the basis functions using the infinite element method. Our methods can handle singularities efficiently and are not sensitive to the large contrast.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Overview

Solving the singularly perturbed convection-diffusion equation accurately has long been a challenge for numerical analysts. There is a large literature on the development of specialized numerical methods for these problems, summarized in [73, 63, 50] among others. However, most of the work has been confined to the one-dimensional or very simple two-dimensional cases. Effective methods in one dimension can be derived without too much difficulty, but they are typically extended to two dimensions via operator splitting along coordinate directions and then by the application of the one-dimensional method in each direction. Therefore, the convective term has to be simple and the solution has to be essentially one-dimensional in its behavior in order for these methods to work. For problems with genuinely two-dimensional effects, these methods often perform poorly. In the present work, we discuss effective methods for some nontrivial two-dimensional problems.

The difficulty in obtaining the numerical solution is in large measure due to the dual nature of the convection-diffusion equation. Because the diffusive term is multiplied by a small parameter ϵ ($0 < \epsilon \ll 1$), the convective effect dominates and the equation behaves essentially as a hyperbolic one in a large part of the domain. In a small region, however, diffusion becomes important and the balance between the convective and diffusive effects creates boundary or internal "layers." In these layers, the solution varies rapidly over a short distance. Numerically, this means that a very fine grid is required, at least inside these regions, in order to compute the solution accurately. In particular, the layers typically have the exponential behavior of the form $e^{f(\mathbf{x})/\epsilon^\alpha}$, with some smooth function $f(\mathbf{x})$ and exponent α for ϵ. The thickness of these layers is proportional to ϵ^α. Therefore, as $\epsilon \to 0$, the thickness of the layers decreases accordingly and the number of grid points required is often such that it is infeasible to resolve the layer structure by simply taking smaller and

smaller mesh size.

Many methods have been proposed to deal with this problem. These common methods basically fall into two categories. The first type uses a modification of the underlying mesh. The idea is to put more grid points only where they are needed. For simple problems in which much about the layers is known, optimal meshing and its properties have been developed, as summarized in [63]. For more complicated problems, general adaptive meshing algorithms using *a posteriori* error estimates are needed; see, e.g., [12]. However, these algorithms are complicated and require sophisticated programming. Unlike the methods we consider in the present work, the grid points added for refinement in adaptive meshing algorithms are coupled globally to the grid points of the regular mesh, increasing the size of the matrix problem. The linear system of the discretized equations also lacks the banded structure of the uniform grids. In [36], an example is given in which a mesh refinement for an internal layer problem gives an indefinite matrix even though the differential operator is positive definite. All this makes solving the linear system more difficult and time-consuming as the level of refinement increases. Moreover, a standard grid adaptation in which more mesh points are placed where the solution changes rapidly may not always place transition layers correctly. It has been shown in [21] that to compute these layers correctly, the numerical method may need to include a large number of points close to the boundary where the solution may be flat.

The second class of methods for dealing with the layers is based on a modification of the differential operator. The simplest approach in this category is the upwinding method (see, e.g., [57, 80]), in which differencing schemes are biased in the direction of the convection. In two dimensions, the characteristics of the hyperbolic part of the equation are followed. This method does produce a stable numerical solution, but at the loss of accuracy. In general, it is $\mathcal{O}(h)$ accurate away from the layers but only $\mathcal{O}(1)$ near the layer [53]. Moreover, the error bounds depend on the Sobolev norm of the solution, which is large for small diffusion. Another way of suppressing spurious oscillations is the method of artificial viscosity. In two dimensions, it is the streamline-diffusion in which viscosity is added only in the direction of characteristics [50, 29]. But, again, the fronts are smeared and the accuracy is degraded. The global estimates are also not uniform with respect to ϵ, and a parameter must be

tuned in order to add the right amount of artificial viscosity. A different approach involves adjusting the differencing scheme. Since the layers are typically exponential, the idea is to modify the differencing scheme in such a way that an exponential function is captured exactly rather than a polynomial one. This idea of "exponential fitting" is very old, dating back to 1950s [2], and there have been many variants on this theme [62, 68, 81, 39, 59, 76, 72]. In the finite element setting [26, 20], the finite element space is modified with basis functions that have the exponential behavior. By reflecting the multiscale property of the differential operator, these basis functions can improve the solution dramatically. The difficulty, however, has been that it is hard to construct such basis functions in the genuinely two-dimensional problems. This is the subject of the present work.

We find that there are in general two classes of velocity fields, for which different approaches are needed. In Chapter 2, we study the case in which the correct behavior of the solution can be determined by examining the equation *locally*. We study this case in the context of the enhanced effective diffusivity problem from fluid mechanics [7, 5, 45]. This provides an interesting physical background for examining the performance of our method. Specifically, we study the transport properties of the highly oscillatory but periodic cellular velocity field. The fluid in this flow is rotating in each of the small cells of size δ, with opposite directions in adjacent cells. To this problem, we apply the multiscale finite element method originally developed for elliptic problems [44, 43], and obtain the correct scaling of the overall diffusivity property of the flow. In this method, the support of each basis function is larger than the small scale of the problem, and we obtain the basis by solving the associated homogeneous equation locally in place of the standard polynomial basis. These basis functions then contain the oscillations from the small scales. Thus, rather than resolving the small scales of the coefficient in a globally coupled fashion, we resolve them within each basis function. The finite element formulation then transfers the correct averaged effect to the coarse grid. Thus, we break up an otherwise prohibitively large problem into smaller, more manageable pieces. Results from homogenization theory [49, 16, 77] are introduced in Section 1.4 to be used in the analysis of this method. In Section 2.4, we carry out the asymptotic error analysis for the method to show that the relative error is of order δ/h, and does not depend on the small scale ϵ. This estimate shows that for a fixed grid h, the method converges in the limit of $\delta \to 0$. This estimate is sharp, as verified by numerical

computations. In contrast, the standard finite element method has δ on the denominator and the method fails to converge unless we have a very fine mesh so that $h \ll \delta$.

For this analysis, we assume that the details within each basis functions are well-resolved. While the construction of each basis is independent and therefore easily parallelized, it can still be expensive if ϵ is so small that a large number of mesh points are required inside an element. In Section 2.5, we explore the possibility of using an asymptotic expansion, in place of a direct computation. While we find that we can obtain a good approximation this way for the cellular flow problem at hand, capturing two-dimensional behavior with asymptotics seems too difficult in general.

The problems we consider in Chapter 2 are manageable because the characteristic length scale that governs the local behavior is small. The information contained in the support of each element is sufficient to determine the local behavior correctly. This allows us to break up the problem into smaller pieces and still capture the correct overall solution. In Chapter 3, we examine the more difficult case in which the characteristic length scale is much larger than the element size. This means we cannot obtain useful basis functions by solving the equation locally; the support of each basis function is too small to contain the *nonlocal* information. Thus, in order to come up with the basis functions of the correct local behavior, we must have some knowledge of the solution structure *a priori*. When this information is not available, one must then resort to conventional methods; but for many problems, enough information is available or can be obtained to aid us in designing a more effective method. For the random velocity field we consider, for instance, theoretical understanding of the flow field based on percolation theory [33] and the variational principle introduced in [31] can provide the necessary information. This variational principle is nonlocal and a test function must be evaluated on the whole domain in order to compute the energy integral. While the test function that minimizes the integral has a complicated layer structure and is difficult to find, we can construct an artificial test function that captures the essential behavior of the layers and use this to extract the location and the thickness of boundary layer. The information obtained this way is interesting by itself, because it gives an estimate of the effective diffusivity property of the flow introduced in Chapter 2. Once we have the information regarding the layers, we examine how to use this to construct useful basis

functions. We first consider a one-dimensional problem with a turning point to get insights into how to construct such basis functions in two dimensions. In both cases, we find it necessary to embed the exponential behavior into separate basis functions in order to handle internal layers. These basis functions are then placed along the streamlines on which the layers occur, while the bilinear basis functions of the underlying uniform grid are still kept for the smooth part of the solution. When the exponential bases are incorporated correctly into the finite element formulation, we find that the numerical solution can be obtained from a relatively coarse grid and that the error is reduced substantially. In particular, the exponential basis functions adapt well with decreasing ϵ.

In the last chapter, we study an elliptic equation with large variations in the coefficient [52, 55, 17, 18]. This "high contrast" situation is common, since the coefficient often represents a property of a medium which may differ drastically in different regions. When the coefficient has extremely large variation or even discontinuities, singular behavior results and conventional numerical methods do not perform well [10, 58]. For certain singularities, the standard finite element method can be arbitrarily slow in its convergence [75]. To remedy the situation, we present another variation on the finite element method approach we have employed in Chapter 2 and Chapter 3. The goal again is to build special basis functions that can capture the correct local behavior. This elliptic problem is similar to the convection-diffusion problem of Chapter 2 in that the singularity is essentially local. But an additional difficulty is that we may not be able to solve the equation locally: some singularities, such as the "checkerboard" pattern, are nearly impossible to resolve using any conventional scheme. In some instances, the form of the singular behavior can be obtained, for example, through the solution of an eigenvalue problem [79, 66]; numerical methods have been developed using this information. But in general, the nature of the singular behavior is unknown and we need another means of obtaining the local solution. The solution we present here is the construction of the local basis using the infinite element method [82]. This elegant technique takes advantage of the similarity of its special grid structure and has the effect of having infinite resolution at the singularity. The number of unknowns in the solution process is proportional to the number of points on the element boundary, i.e., given m mesh points on each side of an element, the matrix problems involved are for $4m$ unknowns, not for m^2 unknowns as in most other methods. We also derive a way of com-

puting the stiffness matrix efficiently. We examine the case of a corner and checkerboard singularities [75] and find that the method performs very well. As the contrast increases, the singularity gets more localized and even for a moderate mesh size in the infinite element, the method is not sensitive to the high contrast.

In summary, we pursue the numerical solution of singularly perturbed or degenerate elliptic problems within the finite element framework. When the nature of the problem is such that we can capture the multiscale property of the differential operator locally, we construct special finite element basis functions by solving the equation directly within the elements. In particular, we employ the multiscale finite element method for highly oscillatory problems and the infinite element technique for the singularities that cannot be resolved by other means. When the layer structure is determined nonlocally, we make use of the variational principle or matched asymptotics results to design the basis functions with correct features. We verify the effectiveness of the schemes with numerical experiments.

1.2 The Convection-Diffusion Equation

1.2.1 Motivation

Many phenomena in science and engineering are transport processes. Some quantity of interest gets transported by the diffusive motion of particles and by the convective motion of the underlying velocity field. For example, the transported quantity $u(\mathbf{x}, t)$ may be the temperature or the concentration of a pollutant. The partial differential equation that describes these effects is the convection-diffusion equation

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{b}(\mathbf{x}) \cdot \nabla u(\mathbf{x}, t) - \epsilon \Delta u(\mathbf{x}, t) = f(\mathbf{x}). \tag{1.1}$$

The term "advection" is sometimes used in place of convection interchangeably. We assume that the velocity field $\mathbf{b}(\mathbf{x})$ is given in the problem. One may, for example, solve the Navier-Stokes equations coupled to this equation to obtain this velocity field. ϵ $(0 < \epsilon \ll 1)$ is the molecular diffusivity characterizing the Brownian motion; f is the forcing on the

system. The convection term is sometimes written in a more general form as $\nabla \cdot (\mathbf{b}u)$. But most velocity fields in which we are interested come from streamfunctions and hence are incompressible, $\nabla \cdot \mathbf{b} = 0$.

Examples of this equation are found in many areas of science and engineering. It is the linearized version of the Navier-Stokes equations with $1/Re$ being the small parameter. Here Re is the Reynolds number which is very large in many important problems. It is also the vorticity equation for the incompressible Navier-Stokes equations in two dimensions with viscosity ϵ. Transport of heat in Rayleigh-Bénard convection [74] and transport of magnetic field in the kinematic dynamo problem [65, 78, 25] are also described by (1.1). In semiconductor device simulation, it is the widely-used "drift-diffusion" model [69] with u being the electric potential.

1.2.2 The Singularly Perturbed Case

An equation is said to be singularly perturbed when the term with the highest derivative is multiplied by a small parameter. In the present case, we also refer to it as the convection-dominated case, since $|\mathbf{b}|/\epsilon$ is large. In fluid mechanics, it is also referred to as the high Peclet number problem. The Peclet number Pe is the ratio between the convective and the diffusive effects.

Equation (1.1) can be written in elliptic form for any positive viscosity ϵ if we can find a skew-symmetric matrix $\mathbf{H} = (H_{ij})$ such that $\nabla \cdot \mathbf{H} = -\mathbf{b}$. This is the case if $\mathbf{b}$ is incompressible and has mean zero. In two dimensions, if $\mathbf{b}$ is obtained through a streamfunction ψ, that is, $\mathbf{b} = \nabla^{\perp}\psi = (-\psi_y, \psi_x)$, we simply have

$$\mathbf{H} = \begin{pmatrix} 0 & -\psi \\ \psi & 0 \end{pmatrix}. \tag{1.2}$$

Now, the equation (1.1) can be written in the following divergence form:

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} - \nabla \cdot (\epsilon \mathbf{I} + \mathbf{H}) \nabla u(\mathbf{x}, t) = f(\mathbf{x}). \tag{1.3}$$

As mentioned in the overview, the limiting equation with $\epsilon \to 0$ is hyperbolic, and the equation qualitatively does not have many features that we would expect from an elliptic equation. In particular, solutions of (1.3) may have sharp boundary and internal layers. The boundary layer theory for the singularly perturbed problem has been studied for a long time, and there are some analytical techniques available, such as the matched asymptotics and multiple scale expansions [15, 40, 54, 67, 41]. However, except for simple problems, the analytical approach is limited and we must turn to numerical schemes. Unfortunately, the standard numerical schemes have difficulties of their own. In the present work, we improve the numerical methods by combining them with some analytical understanding of the layer structure.

An important characteristic of (1.1) is that the operator is not self-adjoint and we do not have the nice properties of the Sturm-Liouville type problems. The stiffness matrix of the finite element formulation is therefore nonsymmetric, requiring a different set of linear system solvers from the most commonly used ones. The matrix problem may already be ill-conditioned because of small ϵ. Also, the lack of self-adjointness creates additional difficulties when we try to formulate a variational principle in Chapter 3.

1.2.3 The Nonlocal Behavior

In Section 1.1, we referred to the fact that there are no simple and efficient methods for general two-dimensional problems. We clarify one reason for this here. A common extension of successful one-dimensional methods to two-dimensional situations is the tensor product approach, in which the one-dimensional solution is used in each direction, i.e., $\phi_{ij}(x,y) = \phi_i(x)\phi_j(y)$. This works well for a small class of problems whose behavior is essentially one-dimensional, e.g., when $\mathbf{b}(\mathbf{x}) = (b_1(x), b_2(y))$ with $b_1(x) > b_0 > 0$ and $b_2(y) > b_0 > 0$ for some constant b_0 [68, 81, 39]. Other methods that work well in genuinely two-dimensional problems impose stringent conditions on the coefficients or require that much information is given about the layer structure in advance [63]. For example, many methods require that the velocity field does not have a turning point where the coefficient switches the sign.

There is a fundamental reason for the difficulty of constructing efficient two-dimensional

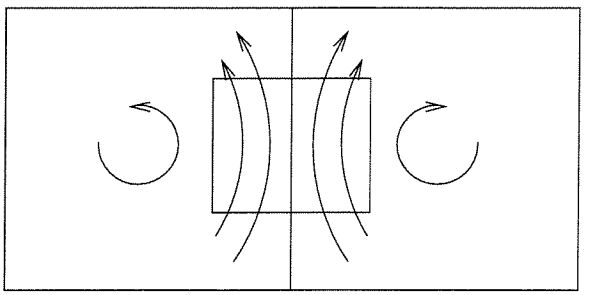

Figure 1.1: A nonlocal effect in the cell problem: in this periodic cellular flow given by the streamfunction $\psi = \sin(\pi x)\sin(\pi y)$ for unit size cells, an internal layer is formed along the vertical boundary at the center; however, this cannot be predicted if the velocity field is sampled only in the box shown.

numerical methods: the layer structure, in particular its location and thickness, is *nonlocal.* We illustrate this in Figure 1.1. Suppose that the fluid is rotating in opposite directions in the adjacent cells, as shown in the picture. Such a velocity field results from the streamfunction $\psi = \sin(\pi x)\sin(\pi y)$ if the cells are of unit size. Locally, near the center boundary, the velocity field is pushing the fluid upward; if we were solving the equation in that small box shown, we would expect the layer to be horizontal along the top. However, the correct solution in this case has vertical layers along the center boundary, determined by the global feature of the velocity field and boundary conditions. This illustrates that unless the sampled region for determining the solution behavior is larger than some characteristic length scale of the problem, we cannot obtain a general method for dealing with layers. Either the sampled region must be large enough or some *a priori* knowledge concerning the structure of the global solution is necessary.

1.3 Formulation of the Problem

We need to introduce some basic spaces before stating the problem in the weak form for the finite element method. Let $\Omega \subset \mathcal{R}^2$ be a bounded domain with a Lipschitz continuous boundary $\partial\Omega$. We will use $L^2(\Omega)$ based Sobolev space $H^k(\Omega)$, which is the space of all functions $u \in L^2(\Omega)$ whose derivatives $D^\alpha u$ (in the sense of distributions) of order $|\alpha| \leq k$

are also in $L^2(\Omega)$. This space is equipped with norms and seminorms

$$\|u\|_{k,\Omega} = \left(\int_\Omega \sum_{|\alpha| \le k} |D^\alpha u|^2 \right)^{1/2},$$

$$|u|_{k,\Omega} = \left(\int_\Omega \sum_{|\alpha| = k} |D^\alpha u|^2 \right)^{1/2},$$

$$\|u\|_{\infty,\Omega} = \operatorname*{ess\,sup}_{x \in \Omega} |u(x)|.$$

The space $H_0^1(\Omega)$ is the closure of the set $C_0^\infty(\Omega)$ in $H^1(\Omega)$. With the sufficiently smooth boundary $\partial\Omega$, $H_0^1(\Omega)$ is the set of all functions u in $H^1(\Omega)$ such that $u = 0$ on $\partial\Omega$. $H^{-1}(\Omega)$ is the dual space of $H_0^1(\Omega)$, the set of all continuous linear functions on $H_0^1(\Omega)$. $H^{1/2}(\Omega)$ is the trace on $\partial\Omega$ of all functions in $H^1(\Omega)$ with the norm $\|v\|_{1/2,\partial\Omega} = \inf \|u\|_{1,\Omega}$ where the infimum is taken over all $u \in H^1(\Omega)$ with trace v. C denotes a generic positive constant independent of small parameters unless otherwise stated and $C + C = C$ and $C \cdot C = C$. We say that a quantity p is $\mathcal{O}(q)$ when $|p| \le Cq$ for all q sufficiently small. We also use the Einstein summation convention, in which we sum over repeated indices.

1.3.1 Weak Formulation

The steady-state problem is

$$-\epsilon \Delta u + \mathbf{b}(\mathbf{x}) \cdot \nabla u = f \quad \text{in} \quad \Omega \subset \mathcal{R}^2, \tag{1.4}$$

$$u = 0 \quad \text{on} \quad \partial\Omega. \tag{1.5}$$

The weak form is to find $u \in H_0^1$ such that

$$a(u, v) = f(v), \quad \forall v \in H_0^1, \tag{1.6}$$

where

$$a(u,v) = \epsilon \int_\Omega \nabla u \cdot \nabla v \, d\mathbf{x} + \int_\Omega \mathbf{b} \cdot \nabla u \, v \, d\mathbf{x},$$
$$f(v) = \int_\Omega f v \, d\mathbf{x}.$$

We assume that the coefficients satisfy the uniform ellipticity condition $c_1|\xi_i|^2 \leq a_{ij}\xi_i\xi_j \leq c_2|\xi_i|^2$ $(0 < c_1 < c_2)$ for any $\xi \in \mathcal{R}^2$, where $a_{ij} = \epsilon\delta_{ij} + H_{ij}$ of (1.3). We assume that $a_{ij}, b_i \in L^\infty(\Omega)$ and $f \in H^{-1}(\Omega)$. We only consider the homogeneous Dirichlet boundary condition for clarity, but other boundary conditions can be easily incorporated. A nonhomogeneous Dirichlet condition can also be easily translated into the forcing term; Neumann boundary condition is automatically enforced by the variational form if Dirichlet condition is not imposed. We fix the domain to be $\Omega = (0,1) \times (0,1) \subset \mathcal{R}^2$ for computation.

1.3.2 Existence and Uniqueness

We need the following theorem for the existence and uniqueness of the solution in the weak form.

Lemma 1.1 (Lax-Milgram Lemma) *Let V be a Hilbert space with some norm $\|\cdot\|_V$, $a(\cdot,\cdot)$ be a bilinear functional, and $f(\cdot)$ be a linear functional such that*

(1) $a(v,v) \geq \alpha \|v\|_V^2, \quad \forall v \in V$ (coercivity)

(2) $|a(u,v)| \leq \beta \|u\|_V \|v\|_V, \quad \forall u, v \in V$ (continuity)

(3) $|f(v)| \leq \gamma \|v\|_V$

for some positive constants α, β, and γ. Then there exists a unique solution such that $a(u,v) = f(v)$, $\forall v \in V$.

Proof. The proof can be found in many books, e.g., [20]. □

Now we can easily obtain the existence and uniqueness for (1.6) by applying the Lax-Milgram Theorem with $V = H_0^1(\Omega)$ and $\|\cdot\|_V = \|\cdot\|_{H^1}$. Conditions (2) and (3) are easily verified; for (1), we have

$$\begin{aligned} a(v,v) &= \epsilon \int_\Omega |\nabla v|^2 + \int_\Omega \mathbf{b} \cdot \frac{1}{2} \nabla (v^2) \\ &= \epsilon \int_\Omega |\nabla v|^2 + \int_\Omega \left(-\frac{1}{2} \nabla \cdot \mathbf{b} \right) |v|^2. \end{aligned}$$

Therefore, we need to have $-1/2\, \nabla \cdot \mathbf{b} \geq 0$. Then, with the Poincaré-Friedrichs inequality [20]

$$\|u\|_{L^2}^2 \leq c_p \, |u|_{H^1}^2 \quad \forall u \in H_0^1,$$

we satisfy condition (1). The incompressible velocity fields we consider satisfy this by $\nabla \cdot \mathbf{b} = 0$.

1.3.3 The Finite Element Method

In the finite element method (FEM) [26, 71], we apply the differential operator exactly but restrict the weak form (1.6) to a finite dimensional subspace of $H_0^1(\Omega)$: Find $u^h \in V^h$ such that

$$a(u^h, v^h) = f(v^h), \quad \forall v^h \in W^h, \tag{1.7}$$

where $V^h = \text{span}\{\phi_1, \ldots, \phi_N\}$ and $W^h = \text{span}\{\psi_1, \ldots, \psi_N\}$ with *trial* and *test* functions ϕ_i and ψ_i, respectively. V^h is called the trial function space and W^h the test function space. When $V^h = W^h$ ($\phi_i = \psi_i$), we refer to the method as the Galerkin method; when $V^h \neq W^h$, we refer to it as the Petrov-Galerkin method.

We choose $v = v^h$ in (1.6) and subtracting (1.7) from (1.6), we have

$$a(u - u^h, v^h) = 0, \tag{1.8}$$

an orthogonality property that becomes very useful in the error analysis. Applying coer-

civity with $v = u - u^h$, orthogonality, and continuity, we have

$$\begin{aligned}\|u - u^h\|_{1,\Omega} &\leq \frac{1}{\alpha} a(u - u^h, u - u^h) \\ &\leq \frac{1}{\alpha} a(u - u^h, u - v^h) \\ &\leq \frac{\beta}{\alpha} \|u - u^h\|_{1,\Omega} \|u - v^h\|_{1,\Omega}.\end{aligned}$$

So we have the following lemma [20]:

Lemma 1.2 (Cèa's Lemma) *Let u be the weak solution of (1.6) and the approximation u^h of (1.7). Then we have*

$$\|u - u^h\|_{1,\Omega} \leq \frac{\beta}{\alpha} \min_{v^h \in V^h} \|u - v^h\|_{1,\Omega}, \quad \forall v^h \in V^h. \tag{1.9}$$

This guarantees that the weak form picks out the optimal v^h from V^h. This version of the lemma, however, is not too helpful for the singularly perturbed case. In this case, we can explicitly compute the constants and get $\beta/\alpha = (1 + c_p^2)^{1/2}(1 + Pe^2)^{1/2}$, where c_p is the constant in the Poincaré-Friedrichs inequality and Pe is the Peclet number $\|\mathbf{b}\|_{L^\infty}/\epsilon$. For our computational domain $(0,1) \times (0,1)$, $c_p = 1/4$. This is the standard step that reduces the finite element method to an approximation theory problem, but for large Pe cases in which we are interested, the $1/\epsilon$ factor makes the estimate less useful.

Finally, we note the standard H_0^1 estimates for interpolants:

$$\begin{aligned}\|u - u^h\|_0 &\leq C\, h^k |u|_k, \\ \|u - u^h\|_1 &\leq C\, h^k |u|_{k+1}.\end{aligned}$$

1.4 Homogenization

In the following chapters, we consider problems with another small scale δ, in addition to the small diffusivity ϵ. This δ is the fine scale characterizing the rapid variation in the

coefficient. In Figure 1.1, for example, the size of the cells may be $\mathcal{O}(\delta)$. In this section, we ignore the issue of boundary and internal layers temporarily and discuss how to deal with this new fine scale.

The rapidly oscillating coefficient is common to many multiple scale problems. If the coefficient were the permeability of porous media, for instance, it would oscillate rapidly. As is the case with ϵ, it may be prohibitively expensive to solve the problems with enough resolution to resolve the fine scale δ. We are often content with the coarse scale behavior of the solution, but we cannot simply ignore the fine scales because the fine scales interact with other scales to affect the coarse grid solution.

Homogenization is a way of extracting an equation for the coarse behavior that takes into account the effect of the small scales. The homogenized or the "effective" equation contains no small scales and is therefore much easier to solve. Physically speaking, it is a method by which a problem in a medium consisting of a large number of periodic cells is approximated by a problem in a homogeneous medium. The classic reference for this is [16], while a more readable account is given in [49]. Another book from a more physical viewpoint is [77]. Many applications in science and engineering are described in [61]. Here, we review the basic homogenization theory that will be important in the analysis later.

We consider the divergence form of the equation,

$$-\nabla \cdot (a\,(x/\delta)\,\nabla u) = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega, \tag{1.10}$$

with $u \in H_0^1(\Omega)$ and $f \in H^{-1}(\Omega)$. The convection-diffusion equation (1.4) can be written in this form, with $a_{ij} = \epsilon\delta_{ij} + H_{ij}$. According to the homogenization theory, the solution u has the following property of convergence,

$$u \to u_0 \qquad \text{weakly in } H_0^1(\Omega), \tag{1.11}$$

$$a\,(x/\delta)\,\nabla u \to a^*\nabla u_0 \qquad \text{weakly in } L^2(\Omega), \tag{1.12}$$

as $\delta \to 0$, where u_0 satisfies the homogenized equation

$$-\nabla \cdot a^* \nabla u_0 = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega. \tag{1.13}$$

The homogenized coefficient a^* does not contain the small scale δ.

One way to derive an expression for the homogenized coefficient a^* is through asymptotic analysis. We introduce the fast variable $y = x/\delta$ and let $a(x, y)$ be periodic in y. We consider x and y to be independent and

$$\frac{d}{dx_i} = \frac{\partial}{\partial x_i} + \frac{1}{\delta}\frac{\partial}{\partial y_i}. \tag{1.14}$$

It is natural [77] to seek the first approximation in the form

$$u(x, y) = u_0(x, y) + \delta u_1(x, y). \tag{1.15}$$

The justification for this expansion is given in [16, 49]. In terms of the separated variables, we have

$$\begin{aligned} L_\delta &= \frac{d}{dx_i}\left(a_{ij}\left(\frac{x}{\delta}\right)\frac{d}{dx_j}\right) \\ &= \left(\frac{\partial}{\partial x_i} + \frac{1}{\delta}\frac{\partial}{\partial y_i}\right) a_{ij}\left(\frac{\partial}{\partial x_i} + \frac{1}{\delta}\frac{\partial}{\partial y_i}\right) \\ &= \frac{1}{\delta^2}A_1 + \frac{1}{\delta}A_2 + A_3, \end{aligned}$$

where

$$\begin{aligned} A_1 &= \frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right), \\ A_2 &= \frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial x_j}\right) + \frac{\partial}{\partial x_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right), \\ A_3 &= a_{ij}(y)\frac{\partial^2}{\partial x_i \partial x_j}. \end{aligned}$$

When we apply this L_δ to the expansion of $u(x,y)$ (1.15), we have

$$L_\delta = \left(\frac{1}{\delta^2}A_1 + \frac{1}{\delta}A_2 + A_3\right)(u_0 + \delta u_1) \tag{1.16}$$

$$= \frac{1}{\delta^2}A_1 + \frac{1}{\delta}\left(A_1u_1 + A_2u_0\right) + \left(A_3u_0 + A_2u_1\right) + \delta A_3u_1. \tag{1.17}$$

We now match the different order terms. The $\mathcal{O}(\delta^{-2})$ term is

$$\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial u_0(x,y)}{\partial y_j}\right) = 0, \tag{1.18}$$

which shows that $u_0(x,y) = u_0(x)$, [16]. The next order gives

$$\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial u_1(x,y)}{\partial y_j}\right) = -\frac{\partial}{\partial y_j}a_{ij}(y)\frac{\partial u_0}{\partial x_j}.$$

This can be considered as a periodic boundary value problem with respect to y. The general solution to this problem is

$$u_1(x,y) = \chi^k(y)\frac{\partial u_0}{\partial x_k}, \tag{1.19}$$

where χ^k is the solution of the "cell problem"

$$\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial \chi^k(y)}{\partial y_j}\right) = -\frac{\partial}{\partial y_i}a_{ik}(y). \tag{1.20}$$

The χ^k function is determined up to a constant; we impose $\langle \chi^k \rangle = 0$ to get a unique solution, where $\langle\cdot\rangle$ denotes the average over one period. With the $\mathcal{O}(\delta^0)$ terms, we take the average with respect to y, which results in the expression for the homogenized coefficient

$$a^*_{ij} = \frac{1}{|Y|}\int_Y a_{ik}\left(\delta_{jk} + \frac{\partial \chi^j}{\partial y_k}\right)d\mathbf{y}, \tag{1.21}$$

where Y denotes the periodic cell of y variable and $|Y|$ is the volume of the cell.

Now, since u_0 satisfies g on the boundary $\partial\Omega$, $u_0 + \delta u_1$ does not. In order to enforce this boundary condition, we need to introduce a "corrector" θ_δ, which satisfies

$$L_\delta\theta_\delta = 0 \quad \text{in} \quad \Omega, \qquad \theta_\delta = -u_1(x, x/\delta) \quad \text{on} \quad \partial\Omega. \tag{1.22}$$

Putting this together, we have

$$u_\delta = u_0 + \delta \chi^k \left(\frac{x}{\delta}\right) \frac{\partial u_0}{\partial x_k} + \delta \theta_\delta. \tag{1.23}$$

This expansion is used in Section 2.4.

1.4.1 A Simple Example

We illustrate this homogenization process with the following simple example:

$$\frac{d}{dx}\left(a\left(\frac{x}{\delta}\right)\frac{du}{dx}\right) = 0, \tag{1.24}$$

where $u \in H_0^1([0,1])$, $u(0) = u(1) = 0$, $a \in L^\infty([0,1])$, and $a(x) \geq a_0 > 0$ for all $x \in [0,1]$.

With $y = x/\delta$, the cell problem (1.20) is

$$\frac{\partial}{\partial y}\left(a(y)\frac{\partial \chi}{\partial y}\right) = -\frac{\partial a(y)}{\partial y}. \tag{1.25}$$

Integrating both sides and dividing by $a(y)$, we get

$$\frac{\partial \chi}{\partial y} = -1 + \frac{c}{a(y)}. \tag{1.26}$$

Since we have $\left\langle \frac{\partial \chi}{\partial y} \right\rangle = 0$ due to periodicity, we find that

$$c = \left\langle a^{-1} \right\rangle^{-1}. \tag{1.27}$$

Using the one-dimensional version of (1.21), we get

$$\begin{aligned} a^* &= \left\langle a(y) + a(y)\frac{\partial \chi}{\partial y} \right\rangle & (1.28)\\ &= \left\langle a(y) + a(y)\left(-1 + \frac{\left\langle a^{-1} \right\rangle^{-1}}{a(y)}\right) \right\rangle & (1.29)\\ &= \left\langle a^{-1} \right\rangle^{-1}. & (1.30) \end{aligned}$$

Thus, the homogenized solution u_0 satisfies

$$\frac{d}{dx}\left(\left\langle a^{-1}\right\rangle^{-1}\frac{du_0}{dx}\right) = 0. \tag{1.31}$$

A naive guess for the homogenized coefficient may be the arithmetic average $\langle a\rangle$. But there can be a large difference between this and the correct coefficient, the harmonic average $\left\langle a^{-1}\right\rangle^{-1}$. If $a(y)$ is the material coefficient that oscillates evenly between the two constant values $a_1 = 1$ and $a_2 = 1000$, then $\langle a\rangle \approx 500$, while $\left\langle a^{-1}\right\rangle^{-1} \approx 2$.

Chapter 2 The Convection-Diffusion Equation with Periodic Coefficients

2.1 Introduction

While equation (1.1) explicitly contains only one small parameter ϵ, another parameter enters when we want to consider the behavior of this equation in the large-domain and long-time limit. Rather than trying to solve the equation in successively large domains for increasingly longer time, we introduce the rescaling parameter δ. We let

$$\mathbf{x}' = \mathbf{x}/\delta \quad \text{and} \quad t' = t/\delta^2, \quad (0 < \delta \ll 1), \tag{2.1}$$

which transforms (1.1) to, after dropping the apostrophe on $\mathbf{x}'$ and t',

$$\frac{\partial u_\delta}{\partial t} + \frac{1}{\delta}\,\mathbf{b}(\mathbf{x}/\delta)\cdot\nabla u_\delta = \epsilon\Delta u_\delta. \tag{2.2}$$

We have also set $f = 0$, since there is no forcing in the problem we consider in this section. The analysis of this equation with periodic coefficients in the $\delta \to 0$ limit involves homogenization, which we reviewed in Section 1.4.

The homogenization theory applied to the convection-diffusion equation is the following, [60, 31]. As $\delta \to 0$, the solution to the rescaled equation u_δ converges to u, which satisfies the constant coefficient equation

$$\frac{\partial u}{\partial t} = \sigma_\epsilon \Delta u. \tag{2.3}$$

The convergence is in L^2,

$$\lim_{\delta\to 0} \sup_{0\le t\le t_0} \int |u(\mathbf{x},t) - u_\delta(\mathbf{x},t)|^2\, d\mathbf{x} = 0 \tag{2.4}$$

for any t_0. Equation (2.3) means that in the large-domain, long-time limit, the overall behavior can be characterized as diffusive, with the convection term making contribution of varying degree depending on the velocity field. How different velocity fields affect this "effective diffusivity" tensor σ_ϵ is of great interest and it is the test problem for the numerical methods we develop. Using homogenization theory, we can obtain an expression for this σ_ϵ [31]:

$$\sigma_\epsilon(\mathbf{e}) = \epsilon \left\langle (\nabla\chi + \mathbf{e}) \cdot (\nabla\chi + \mathbf{e}) \right\rangle, \tag{2.5}$$

where $\langle\cdot\rangle$ denotes averaging over one period and $\mathbf{e}$ is a unit vector. This comes directly from the expression for the homogenized coefficient (1.21). χ is the solution of the cell problem (1.20), which we can rewrite, with $a_{ij} = \epsilon\delta_{ij} + H_{ij}$, also in a vector form as

$$\nabla \cdot [(\epsilon\mathbf{I} + \mathbf{H})(\nabla\chi + \mathbf{e})] = 0, \tag{2.6}$$

with $\mathbf{H}$ as defined in (1.2). This can be simplified by using $\nabla \cdot \mathbf{H} = -\mathbf{b}$ to

$$-\epsilon\Delta\chi + \mathbf{b} \cdot \nabla\chi + \mathbf{b} \cdot \mathbf{e} = 0 \tag{2.7}$$

on the torus. σ_ϵ is generally a nonsymmetric matrix, but for the streamfunction we consider, there is a symmetry of the form $\mathbf{H}(\mathbf{x}) = -\mathbf{H}(-\mathbf{x})$ and this assures that the σ_ϵ tensor is symmetric [31]. Uniform boundedness of the streamfunction ψ is sufficient for the existence of this homogenization [7].

The full solution to (2.2) has fine scales characterized by δ and ϵ, $(0 < \delta \ll 1, 0 < \epsilon \ll 1)$, and conventional numerical methods perform poorly unless we resolve all the small scales. But to do so is prohibitively expensive. In this chapter, we apply the multiscale finite element method [44, 43, 30] to this convection-diffusion problem. After introducing the multiscale FEM in Section 2.2, we demonstrate its effectiveness for the periodic problem by using it to obtain the correct diffusivity scaling for the cellular flow in Section 2.3. In Section 2.4, we study the convergence rate of the multiscale method using asymptotic error analysis. We find that the estimates derived are sharp, as verified by numerical experiments. Finally, in Section 2.5, we examine possible improvements to the method by obtaining the

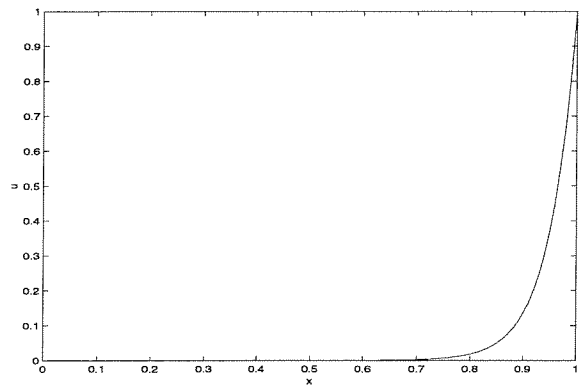

Figure 2.1: 1-D solution with an exponential layer ($a = 1$ and $\epsilon = .05$). There is an exponential layer of thickness $\mathcal{O}(\epsilon)$ at the right boundary.

basis functions using asymptotic methods.

2.1.1 Problems in One Dimension

To gain insights into the behavior of the singularly perturbed equation, we first consider the simple one-dimensional problem,

$$-\epsilon u'' + a(x)u' = 0, \quad x \in [0,1], \tag{2.8}$$

with Dirichlet boundary conditions $u(0) = 0$, $u(1) = 1$. If $a(x)$ is a constant a_0, the exact solution is

$$u(x) = \frac{e^{a_0 x/\epsilon} - 1}{e^{a_0/\epsilon} - 1}, \tag{2.9}$$

which is plotted in Figure 2.1. Near $x = 1$, there is an exponential layer of thickness $\mathcal{O}(\epsilon)$, in which the $\epsilon u''$ term becomes important. Intuitively, we can think of the u' term as the convective effect pushing the "fluid" to the right. When $a < 0$, it gets pushed to the left and the boundary layer is near $x = 0$. If $a(x_0) = 0$ for some $x_0 \in [0,1]$, x_0 is called a turning point. In this case, internal layers can also occur.

Two standard methods for solving (2.8) are the finite difference and the finite element methods. In the simplest FEM, we expand $u(x) = \sum_1^N u_i \phi_i(x)$, where ϕ_i are the "hat"

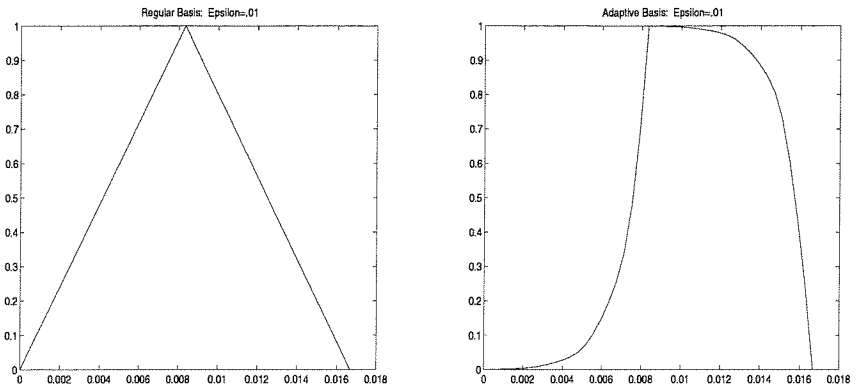

Figure 2.2: Regular and adaptive basis functions ($\epsilon = .01$). The adaptive basis function on the right is obtained by solving the equation locally on each side.

functions

$$\phi_i = \begin{cases} (x - x_{i-1})/h, & x \in [x_{i-1}, x_i] \\ (x_{i+1} - x)/h, & x \in [x_i, x_{i+1}] \end{cases} \tag{2.10}$$

as shown on the left in Figure 2.2. Substituting this in the equation (2.8), we have

$$-\left(1 + \frac{ah}{2\epsilon}\right) u_{i-1} + 2u_i - \left(1 - \frac{ah}{2\epsilon}\right) u_{i+1} = 0, \quad i = 1, \ldots, N-1,$$

where there are $N+1$ grid, or "nodal," points and $h = 1/N$. The finite difference discretization with central differencing gives exactly the same set of equations in this case. With boundary conditions $u_0 = 0$ and $u_N = 1$, the exact solution to this set of equations at the nodal points are

$$u_i = \frac{\gamma^i - 1}{\gamma^N - 1}, \quad \gamma = \frac{1 + ah/(2\epsilon)}{1 - ah/(2\epsilon)}.$$

We immediately see that unless $ah/\epsilon < 2$, oscillations will occur. The ratio ah/ϵ is often called the *mesh Peclet* number. This illustrates the common problem, that the mesh size h needs to be very small when ϵ is very small.

2.1.2 Green's Function Approach

To deal with the present difficulty, many methods have been proposed, as reviewed in Section 1.1. Among the most effective methods are the finite element methods with basis functions that contain the exponential behavior resembling the boundary layer of the solution [73].

It is common [39] to obtain the trial functions by solving the homogeneous equation modified by making the coefficient constant. In one dimension,

$$-\epsilon\phi_i'' + \bar{a}\phi_i' = 0, \quad \text{with } \bar{a} = (a(x_{i-1}) + a(x_i))/2, \quad \text{for } x \in (x_{i-1}, x_i). \tag{2.11}$$

The functions ϕ_i obtained this way have exponential layers, as shown on the right in Figure 2.2. When these functions are summed up with correct weights, they provide a better approximation to the solution than the hat functions do. In fact, if the coefficient is constant, this scheme gives exact nodal values.

While (2.11) was based on a heuristic thinking, we can be more systematic by thinking in terms of the Green's functions. By definition,

$$u(\mathbf{x}) = \int_\Omega G(\mathbf{x}, \mathbf{x}_0) f(\mathbf{x}_0)\, dx_0. \tag{2.12}$$

Choosing $v(\mathbf{x}) = G(\mathbf{x}_0, \mathbf{x})$ in the bilinear form (1.6) and using (2.12), we have

$$a(u, G) = (f, G) = u. \tag{2.13}$$

Note that this is true for any choice of f, which means it is also true for any choice of u. Letting $u \to u - u^h$ in (2.13), we can now write

$$a(u - u^h, G) = u - u^h.$$

Subtracting this from the orthogonality condition $a(u - u^h, v) = 0$ (see equation (1.8)), (1.8)

$$u - u^h = a(u - u^h, G - v), \quad \forall v \in V^h. \tag{2.14}$$

This shows that the error $\|u - u^h\|$ can be minimized by selecting the test function space V^h to contain as much of the Green's functions as possible. Given this reasoning, it makes sense to choose V^h to include functions G_i that satisfy

$$L^* G_i(\mathbf{x}) = -\delta(\mathbf{x} - \mathbf{x}_i)$$

at each mesh point $\mathbf{x}_i$. L^* is the adjoint operator here.

Fortunately, in one dimension, G_i is a linear combination of local Green's functions $g_j(x), x \in (x_{j-1}, x_{j+1})$, that satisfy

$$L^* g_j(x) = 0, \quad g_j(x_{j-1}) = g_j(x_{j+1}) = 0, \quad g_j(x_j) = 1. \tag{2.15}$$

For variable coefficient problems, previously proposed methods suggest solving a local problem with averaged coefficients. However, as we will also do in two dimensions, we solve the variable coefficient problem directly. Then we obtain the following lemma:

Lemma 2.1 *For the one-dimensional equation (2.8) with a variable coefficient convection term, using linear trial functions and solutions to the local homogeneous adjoint equation as test functions in the weak formulation results in exact nodal values, if the integrations are done exactly.*

The proof is in the appendix.

Note that the trial functions can be any basis as long as they are continuous. If there is no forcing, adaptive trial functions and linear test functions also give exact nodal values.

In two dimensions, the global Green's function cannot be expressed as a linear combination of local Green's functions solved in each element. However, the adjoint equation is still

useful and we can see its effect in the following expression:

$$\begin{aligned} a(v,w) &= \epsilon \int_\Omega \nabla v \cdot \nabla w \, d\mathbf{x} + \int_\Omega \mathbf{b} \cdot \nabla v \, w \, d\mathbf{x} \\ &= \epsilon \int_\Omega \nabla v \cdot \nabla w \, d\mathbf{x} - \int_\Omega \mathbf{b} \cdot \nabla w \, v \, d\mathbf{x}, \end{aligned}$$

where the boundary term disappears in the integration by parts because $w \in H_0^1$. If we break up the integral into elements denoted by K and integrate by parts in each element, we get

$$\begin{aligned} a(v,w) &= \epsilon \sum_i \int_K \nabla v \cdot \nabla w \, d\mathbf{x} - \sum_i \int_K \mathbf{b} \cdot \nabla w \, v \, d\mathbf{x} \\ &= \epsilon \sum_i \int_{\partial K} v \nabla w \cdot \mathbf{n} \, ds - \epsilon \sum_i \int_K \Delta w \, v \, d\mathbf{x} - \sum_i \int_K \mathbf{b} \cdot \nabla w \, v \, d\mathbf{x}. \end{aligned}$$

When we get w by solving the adjoint equation

$$-\epsilon \Delta w - \mathbf{b} \cdot \nabla w = 0$$

on each element, the second and third terms disappear. We then let $v = u - u^h$ and use the projection property (1.8) and get

$$\sum_i \int_{\partial K} (u - u^h) \nabla w \cdot \mathbf{n} \, ds = 0.$$

We see that when the test functions w are the solutions of the adjoint problem, there is a projection of the error $u - u^h$ onto the element boundary that becomes zero.

In the next section, we describe a similar method called the multiscale finite element method [44], in which the *trial* functions are the solution of the homogeneous equation and the test functions some continuous function, e.g., the bilinear functions. This method was motivated by a different reasoning for different problems, one to capture sharp layers and one to capture rapid oscillations. But the two methods are similar in that both use the homogeneous equation locally to capture the property of the differential operator more accurately. Numerical comparison is given at the end of Section 2.4.5 for different trial and

test functions.

2.2 The Multiscale Finite Element Method

In [44], the multiscale FEM was introduced for elliptic equations with oscillatory coefficients. The main idea is to resolve the fine scales locally within each element of size $h \gg \delta$ by solving the homogeneous equation with some appropriate boundary conditions. This way, each basis function retains the oscillatory property of the differential operator. The authors prove that when these elements are used to construct the global stiffness matrix, the averaged effects of the rapidly oscillating coefficients are correctly captured. By resolving the fine details inside the elements, problems that are prohibitively expensive with conventional finite element methods are broken down into smaller, manageable parts. Because the elements are larger, the final solution is computed on a coarser grid. However, the large scale features are usually of interest in the first place, and sufficient information is obtained for that purpose.

Formally, the multiscale FEM formulation is the following. We let $\mathbf{K}^h$ be a partition of Ω of triangles or rectangles with diameter less than h. We define a set of basis $\phi^i_{\delta,K}, i = 1, \ldots, d$, ($d = 3$ for triangles and $d = 4$ for rectangles) such that $\phi^i_{\delta,K}$ satisfies

$$L_\delta \phi^i_\delta = 0 \quad \text{in } K. \tag{2.16}$$

At the nodal points $x_j \in K, j = 1, \ldots, d$, we require $\phi^i_{\delta,K}(x_j) = \delta_{ij}$ as usual. The correct boundary conditions for (2.16) would match the global solution, but since we do not know this, we can impose linear boundary conditions for now; this issue is discussed later. We define

$$V^h = \text{span}\{\phi^i_{\delta,K} : i = 1, \ldots, d, \quad K \subset \mathbf{K}^h\} \subset H^1_0(\Omega),$$

and the Galerkin formulation is to find the solution $u^h_\delta \in V^h$ such that

$$a(u^h_\delta, v) = f(v) \quad \forall v \in V^h. \tag{2.17}$$

In order to understand the convergence of this method, the ideas from the homogenization

procedure discussed in Section 1.4 play a critical role. Just as we expand u_δ in (1.15), we can also expand the basis function ϕ_δ as

$$\phi_\delta = \phi_0 + \delta\chi^k\left(\frac{x}{\delta}\right)\frac{\partial\phi_0}{\partial x_k} + \delta\theta^\delta, \tag{2.18}$$

with ϕ_0^i, χ^k, and θ^δ defined similarly to (1.13), (1.20), and (1.22). Since ϕ_0 is a smooth function, we see that the oscillation in ϕ_δ comes from that of χ, which solves the cell problem (2.7). The error analysis is based on the fact that ϕ_δ and u_δ satisfy the same operator and therefore the two expansions (1.23) and (2.18) match, away from the small region near the boundary.

We see that the boundary conditions are an important issue if the two expansions were to match better near the boundary. While we assumed linear boundary condition above, another possibility is to solve the one-dimensional version of the equation along each edge, as we do for the problems in this chapter. This usually results in some improvement in error. For the very small ϵ's, care must be taken: the layers inside the basis may have thickness of $\mathcal{O}(\delta\sqrt{\epsilon})$, but the reduced equation on the boundary may give layers of thickness $\mathcal{O}(\delta\epsilon)$. In that case, the mesh spacing that resolves the layers inside the basis may not resolve the layers on the boundary. The case in which solving the one-dimensional equation works the best is in Chapter 4, when coefficients have discontinuities. For many elliptic problems, the best solution turns out to be the oversampling method introduced in [43]. In that case, the effect of a wrong boundary condition is restricted to a $\mathcal{O}(\delta)$ region near to the boundary. The oversampling idea is then to solve for the basis function on a domain larger than the element and extract the information from the middle of the domain. This reduces the error due to the incorrect boundary conditions. A rigorous analysis has been carried out in [30]. In the convection-diffusion case, the oversampling does not always work because the effect of the boundary condition may not be confined to such a small region.

2.3 Computation of the Effective Diffusivity

We apply the multiscale FEM discussed in the previous section to the time-dependent convection-diffusion problem with the periodic velocity field. In particular, we consider the "cellular" flow for which some analytical results exist. The problem we use to demonstrate the effectiveness of the method is the computation of the effective diffusivity property (2.5).

2.3.1 Effective Diffusivity

The "cellular flow" is given by the streamfunction

$$\psi = \frac{1}{4\pi^2} \sin\left(\frac{2\pi x}{\delta}\right) \sin\left(\frac{2\pi y}{\delta}\right), \quad (x, y) \in [0, 1]^2, \tag{2.19}$$

where δ is the rescaling parameter that determines the number of cells. The scaling factor $1/(4\pi^2)$ is used to relate to other studies done on domains of size $[-\pi, \pi]^2$. The resulting velocity field is

$$\mathbf{b} = \nabla^{\perp}\psi = (-\psi_y, \psi_x) \tag{2.20}$$

$$= \left(\frac{1}{2\pi\delta} \sin\left(\frac{2\pi x}{\delta}\right) \cos\left(\frac{2\pi y}{\delta}\right), \frac{1}{2\pi\delta} \cos\left(\frac{2\pi x}{\delta}\right) \sin\left(\frac{2\pi y}{\delta}\right)\right). \tag{2.21}$$

This streamfunction with $\delta = 1$ is plotted in Figure 2.3. The velocity vector is tangent to the streamlines shown.

A well-known result for this velocity field, described for example in [25, 24, 74], concerns the long-time, large-distance diffusive behavior: the "effective diffusivity" σ_ϵ in the limit $\epsilon \to 0$, scales as

$$\sigma_\epsilon \sim \sqrt{\epsilon}. \tag{2.22}$$

In the absence of convection, the effective diffusivity is simply ϵ. It is clear that convection can only increase the overall diffusivity, as it carries particles along the streamlines faster than diffusion does. This can be seen clearly when we expand the expression for σ_ϵ in (2.5)

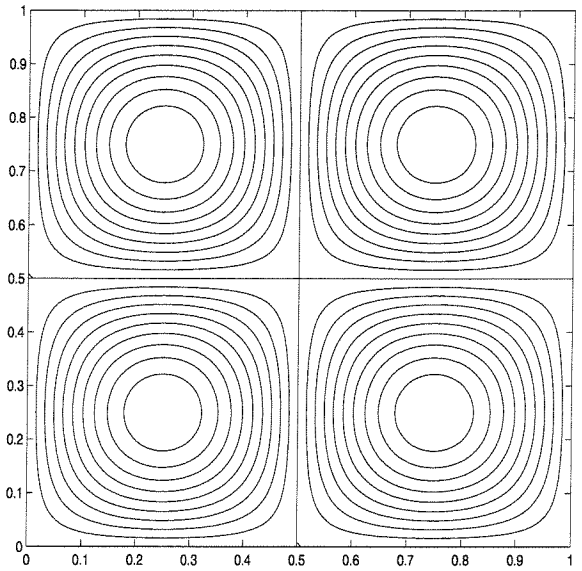

Figure 2.3: A streamfunction for the cellular flow ($\delta = 1$).

and simplify to get another expression,

$$\sigma_\epsilon = \epsilon + \epsilon \langle \nabla\chi \cdot \nabla\chi \rangle , \tag{2.23}$$

where the second term is positive.

The result (2.22) can be understood through a simple scaling argument. The important feature in the solution is the formation of boundary layers near separatrices. It is these boundary layers that characterize the transport properties of particles. To determine the width of the layer along the separatrices, we set $\delta = 1$ and balance the diffusive flux across the layer with the convective flux along the layer. Comparing the time scales,

$$\frac{w^2}{\epsilon} \sim \frac{l}{u_0},$$

where w is the width of the layer, l is the size of the cell, and u_0 is the magnitude of the velocity. Since l and u_0 are of $\mathcal{O}(1)$, we conclude that $w \sim \sqrt{\epsilon}$. Then we can use (2.5). $w \sim \sqrt{\epsilon}$ means $\nabla\chi \sim 1/\sqrt{\epsilon}$ and substituting this in (2.5) and integrating over the width of the layer immediately gives (2.22).

In general, σ_ϵ is a tensor. It is the asymptotic rate of mean square displacement with

different diffusivity depending on the direction. In the present case, however, the effective diffusivity is isotropic. This allows us to introduce a simple and intuitive definition, also useful for computations. With a slight modification from [31], we can measure the mean square displacement by

$$\sigma_\epsilon = \lim_{t\to\infty} \frac{1}{4t} \iint (x^2 + y^2)\, u(x, y, t)\, dx\, dy, \tag{2.24}$$

with the delta function at the origin as the initial function. We have inserted the factor 4 so that when there is no convection term, we get $\sigma_\epsilon = \epsilon$, which we can verify by putting in the Green's function for the heat equation into (2.24).

2.3.2 Numerical Results

In order to test the multiscale FEM, we compute the equation (2.2) in time, starting with a regularized delta function. An example of a multiscale basis function obtained by solving (2.16) is shown in Figure 2.4, to illustrate the fine details within the basis. The basis function has the same layer structure as we would find in the global solution. The boundary conditions given are 1 at the lower left corner and 0 at the other corners; along the edges, the reduced one-dimensional equations are solved. Layers are strong at the lower left region because of this boundary condition.

Discretization

In the weak formulation, the time-dependent problem is

$$(u_t, v) + (b_1 u_x, v) + (b_2 u_y, v) = \epsilon\left((u_x, v_x) + (u_y, v_y)\right).$$

We define a uniform mesh of $(n+1) \times (n+1)$ grid points (ih, jh), $i, j = 0, \ldots, n$, where $h = 1/n$. Labeling the $P = (n+1)^2$ basis functions with a single index, we let $u(x, y, t) =$

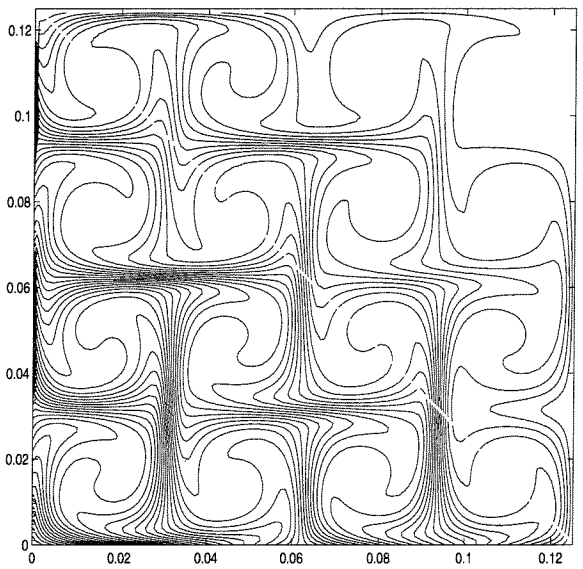

Figure 2.4: A contour plot of the multiscale basis function, with 1 at the lower left corner and 0 at the other nodes. In this example, $\delta/h = .5$, where h is the mesh size.

$\sum_{j=1}^{P} \xi_j(t)\phi_j(x,y)$ and $v = \phi_i$ and get

$$\sum_{j=1}^{P} \xi_j'(t)\,(\phi_j, \phi_i) + \sum_{j=1}^{P} \xi_j(t)\,(b_1\phi_{jx}, \phi_i) + \sum_{j=1}^{P} \xi_j(t)\,(b_2\phi_{jy}, \phi_i)$$
$$= \epsilon \sum_{j=1}^{P} \xi_j(t)\left[(\phi_{jx}, \phi_{ix}) + (\phi_{jy}, \phi_{iy})\right].$$

A good discretization scheme for this problem is second order Adams-Bashforth on the advection term and Crank-Nicholson on the diffusion term. With this, we have

$$\sum_{j=1}^{P} \left\{ (\phi_j, \phi_i) - \frac{\epsilon\Delta t}{2}\left[(\phi_{jx}, \phi_{ix}) + (\phi_{jy}, \phi_{iy})\right] \right\} \xi^{n+1} =$$
$$\sum_{j=1}^{P} \left\{ (\phi_j, \phi_i) - \frac{3}{2}\Delta t\left[(b_1\phi_{jx}, \phi_i) + (b_2\phi_{jy}, \phi_i)\right] + \frac{\epsilon\Delta t}{2}\left[(\phi_{jx}, \phi_{ix}) + (\phi_{jy}, \phi_{iy})\right] \right\} \xi^{n}$$
$$+ \sum_{j=1}^{P} \frac{\Delta t}{2}\left[(b_1\phi_{jx}, \phi_i) + (b_2\phi_{jy}, \phi_i)\right] \xi^{n-1}.$$

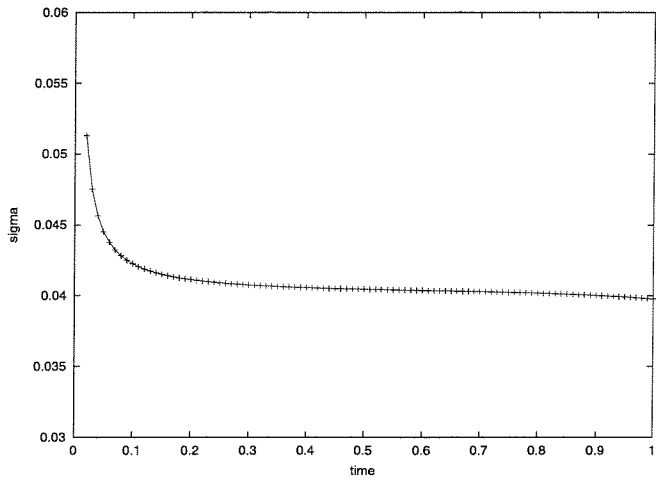

Figure 2.5: Computation of σ_ϵ over time ($\epsilon = .01$).

We can approximate the initial condition with a Gaussian or another smooth function such as

$$u(x,y) = \begin{cases} \frac{1}{16}(\sin(4\pi r + 1/8) + 1)^2, & r < .25 \\ 0, & r > .25 \end{cases},$$

where r is the radius.

We restrict $\mathcal{R}^2$ to a finite domain and then scale it to the computational domain of the unit square using δ; we take small Δt and compute up to some time T, depending on ϵ and δ. When σ_ϵ starts to decrease noticeably, we know that the fluid front has reached the boundary but is prevented from moving further. In Figure 2.5, we see how σ_ϵ evolves in time, before this happens.

We compute the scaling by comparing the values of σ_ϵ at a fixed time for different ϵ. In Table 2.1, we plot the rate at which σ_ϵ in (2.24) is changing. Given the scaling (2.22), we expect σ_ϵ to decrease by a factor of $1/\sqrt{2} \approx .7071$ when ϵ is halved. We see in the table that when $\delta = 1$, σ_ϵ is nearly halved, meaning there is no enhancement. However, when $\delta = .05$, the ratio is close to the predicted value, as ϵ gets small. If the conventional bilinear elements are used, this behavior is missed completely.

We note that while we compute the effective diffusivity here, that is not the main objective of the method. The objective is to compute the transient state correctly. The fact that we

ϵ	σ_ϵ	ratio	σ_ϵ	ratio	σ_ϵ	ratio
	$\delta = 1$		$\delta = .1$		$\delta = .05$	
.04	0.03252		0.03265		0.03275	
.02	0.02092	.6432	0.02316	.7093	0.02355	.7191
.01	0.01141	.5454	0.01643	.7094	0.01665	.7070
.005	0.00619	.5425	0.01300	.7912	0.01233	.7405
.0025	0.00344	.5557	0.01046	.8046	0.00872	.7072

Table 2.1: The diffusivity scaling for the cellular flow ($n = 16, m = 32$).

get the correct scaling is just a consequence of obtaining the correct solution.

Multigrid Solver

One of the major difficulties for the singularly perturbed problems is solving the linear system. For small ϵ, the linear system may be indefinite and often ill-conditioned. This means some iterative techniques would not converge. It was shown in [36] that the positive definiteness of the continuous differential operator is not always mirrored by some discretization schemes, even if the matrix is diagonally dominant. As a result, the standard multigrid algorithm, which is an $\mathcal{O}(N)$ algorithm where N is the total number of unknowns, often performs poorly. Fortunately, there is a modified multigrid algorithm developed specifically for convection-dominated problems [28]. It uses matrix-dependent prolongation and restriction operators that account for the character of the equation. We have found that this algorithm is very efficient and robust, usually converging under 10 or 20 iterations to the residual of 10^{-8}. Iterative methods such as BiCGSTAB or GMRES [13] for non-symmetric matrices also work, but the number of iterations required may be large as ϵ gets small.

Parallel Efficiency

A major advantage of the multiscale FEM algorithm is its parallel efficiency. The construction of one basis function is independent from that of any other; the elements can be divided evenly among all the processors first, and then collected to a "master" processor

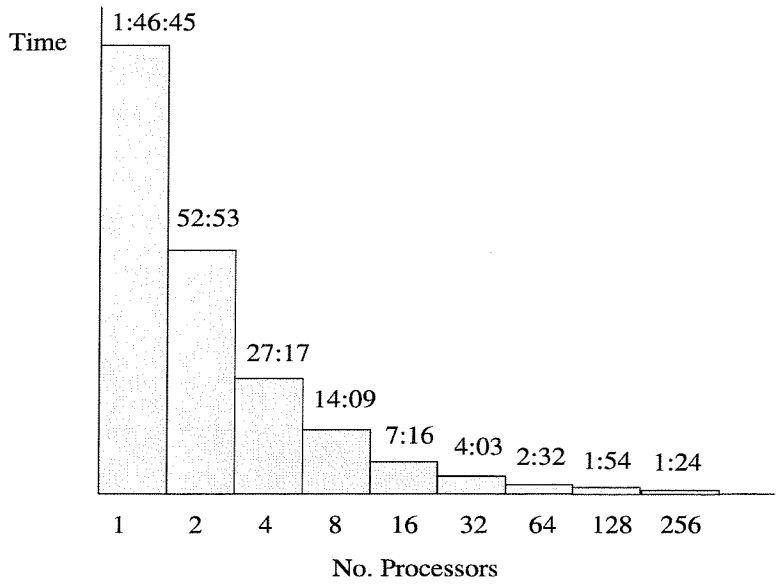

Figure 2.6: Parallel efficiency of the multiscale FEM ($n = 32, m = 64$).

that performs the final calculations. In Figure 2.6, we see that the algorithm is almost perfectly parallelized: as the number of the processors doubles, the computing time is halved. The time is not exactly halved since the master process has to do the final step after the global stiffness matrix is computed. Only when the number of processors is very large is the efficiency degraded, as the communication time amongst the processors takes a significant portion of the total time.

There is also a substantial saving in the memory required. We let n be the number of multiscale elements within each direction and let m be the number of grid points in each element. Then, the total memory required is $\mathcal{O}(n^2 + m^2)$, since $\mathcal{O}(m^2)$ operations for each element can be done in sequence. If the same resolution were to be achieved in direct simulation, $\mathcal{O}\left((nm)^2\right)$ would be required. If $n = 32, m = 32$, for example, there is a factor of 1000 saving in memory.

2.4 Asymptotic Error Analysis

We study the error of the multiscale method in this section. In the genuinely elliptic problem considered before in [44, 30], there is one parameter δ, which characterizes the

rapid oscillation. Now there is an additional parameter ϵ for thin layers that must be considered in examining the error terms.

2.4.1 Previous Results

We first discuss the result for the symmetric coefficient, $a_{ij} = a_{ji}$, studied in detail in [44], as it provides the framework for the analysis of this section. When the small scale of the problem δ is resolved by the mesh size h, the error estimate is

$$\|u - u^h\|_{0,\Omega} \leq C \left(\frac{h}{\delta}\right)^2 \|f\|_{0,\Omega}, \qquad (h \ll \delta). \tag{2.25}$$

In this case, the multiscale basis functions do not have oscillations and look similar to the bilinear functions. Thus, the convergence rate resembles that of the standard FEM. Note that we must have h small enough to resolve the δ scale for convergence. The two methods are different for the case of our interest, $h \gg \delta$. In [44], estimates from homogenization theory are used to show that

$$\|u - u^h\|_{1,\Omega} \leq C_1 \left(\frac{\delta}{h}\right)^{1/2} + C_2 h \|f\|_{0,\Omega}, \qquad (h \gg \delta). \tag{2.26}$$

Note that $\delta \ll 1$ is now the numerator of the first term. The method therefore converges as $\delta \to 0$ for a fixed h, unlike in the $h \ll \delta$ case (2.25). To get the L^2 estimate, the Aubin-Nitsche trick [20] is usually employed to gain an extra order in h from the H^1 estimate. This method, however, does not work well for this problem and results in

$$\|u - u^h\|_{0,\Omega} \leq C_1 \left(\frac{\delta}{h}\right)^{1/2} + C_2 h^2 \|f\|_{0,\Omega}, \qquad (h \gg \delta). \tag{2.27}$$

Notice that the $(\delta/h)^{1/2}$ term does not change. We still expect to see the ratio of δ and h, since these are two scales inherent in the problem. But the exponent $1/2$ is not satisfactory, and numerical experiments show that this $(\delta/h)^{1/2}$ estimate is not sharp.

In order to obtain the correct estimate, subtle error cancellations in the discrete problem

need to be examined. With this analysis, it is formally concluded that

$$\|u - u^h\|_{0,\Omega} \leq C_1 \frac{\delta}{h} + C_2\delta + C_3 h^2 \|f\|_{0,\Omega}, \qquad (h \gg \delta). \tag{2.28}$$

The leading order term is now (δ/h) and as long as h is large enough, the error is small, as verified by numerical examples [44].

Discrete Analysis

We review this analysis of the discrete problem [44, 30], so that we can understand the convection-diffusion case.

Using the triangular inequality, standard finite element estimates for bilinear elements, and the regularity and homogenization estimates [44], we can write

$$\begin{aligned} \|u - u^h\|_{0,\Omega} &\leq \|u - u_0\|_{0,\Omega} + \|u_0 - u_0^h\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega} && (2.29) \\ &\leq C_1\delta + C_2 h^2 \|f\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega}, && (2.30) \end{aligned}$$

where u is the solution to the original continuous problem, u^h the numerical approximation to u given by (2.17); u_0 is the solution to the homogenized problem (1.13), u_0^h the numerical approximation to u_0. We would obtain u_0^h by solving the bilinear form (1.6) with the homogenized coefficient.

The problem then is to estimate the $\|u^h - u_0^h\|_{0,\Omega}$ term in (2.30). To examine this term, we introduce the discrete l^2 norm. It is shown in [44] that

$$\|u^h - u_0^h\|_{L^2} \leq C_1 \|u^h - u_0^h\|_{l^2} + C_2\delta, \tag{2.31}$$

where

$$\|u^h - u_0^h\|_{l^2} = \left(\sum_{i \in \mathcal{N}} \left[u^h(\mathbf{x}_i) - u_0^h(\mathbf{x}_i) \right] h^2 \right)^{1/2} \tag{2.32}$$

with $\mathcal{N}$ containing all the nodal points of the mesh.

Equations (2.30) and (2.31) show that $\|u^h - u_0^h\|_{l^2}$, the convergence of u^h to u_0^h at the nodal points, contains the crucial error term.

Let U^h be the vector containing the nodal points u^h. This is the solution to the discrete equation

$$A^h U^h = f^h, \tag{2.33}$$

where A^h and f^h are the global stiffness matrix and the load vector, obtained from (2.17) using $v = \phi_\delta^h$.

Similarly, we have for the homogenized problem,

$$A_0^h U_0^h = f_0^h, \tag{2.34}$$

where we obtain A_0^h and f_0^h from the bilinear form for (1.13) using $v = \phi_0^i$.

Since the basis function ϕ_i can be expanded as (2.18), we can also expand the stiffness matrix and load vector around the homogenized counterparts as

$$A^h = A_0^h + \delta A_1^h + \cdots, \qquad f^h = f_0^h + \delta f_1^h + \cdots. \tag{2.35}$$

A_1^h is assembled from the local stiffness matrix of each element, which we denote by e. By substituting the expansion $\phi_\delta = \phi_0 + \delta\chi^k \partial\phi_0/\partial x_k + \delta\theta^\delta$ (2.18) in the bilinear form (1.6), we find, after some algebra, that

$$A_{1_{kl}}^e = -\int_K \sigma^{ij}\left(\phi_{0,j}^k \theta_{0,i}^l + \phi_{0,j}^l \theta_{0,i}^k\right) dx + \int_K \delta a_\delta^{ij} \theta_{,i}^k \theta_{,j}^l \, dx + \frac{1}{\delta}\int_K \tilde{\sigma}^{ij} \phi_{0,j}^l \phi_{0,i}^k \, dx \tag{2.36}$$

and

$$f_{1_i}^e = -\int_K f\left(\chi^j \phi_{0,j}^i + \theta^i\right) dx, \tag{2.37}$$

where the comma is a shorthand for partial differentiation. Here,

$$\sigma^{ij} = a^{ik}\left(\delta_{jk} + \frac{\partial \chi^k}{\partial y_j}\right) \tag{2.38}$$

and

$$\tilde{\sigma}^{ij} = \sigma^{ij} - a^{ij}_* - \sigma^{kj}\chi^i_{,y_p}. \tag{2.39}$$

From (1.20) and (1.21), we find that $\left\langle \sigma^{ij} \right\rangle$ is the homogenized coefficient a^{ij}_* and $\sigma^{ij}_{,y_i} = 0$. Integrating by parts, we also obtain $\left\langle \sigma^{pj}\chi^i_{,y_p} \right\rangle = 0$ and hence $\left\langle \tilde{\sigma}^{ij} \right\rangle = 0$ [30].

Given the expansions (2.35), we can deduce an expansion in U^h,

$$U^h = U^h_0 + \delta U^h_1 + \delta^2 U^h_2 + \cdots, \tag{2.40}$$

where U^h_i are given by

$$A^h_0 U^h_i = f^h_i - A^h_1 U^h_{i-1}. \tag{2.41}$$

Finally, we get an expression that we can analyze for U^h_1 from this. We write the $i = 1$ case of (2.41) as

$$U^h_1 = G^h_0 f^h_1 - G^h_0 A^h_1 U^h_0, \tag{2.42}$$

where $G^h_0 = (A^h_0)^{-1}$.

Now the whole analysis reduces to estimating the order of U^h_1, since $\|U^h - U^h_0\| \leq \delta\|U^h_1\| + \delta^2\|U_2\|^2 + \cdots$. With $\|U^h_1\|$, (2.30) and (2.31) directly lead to the desired estimate of $\|u - u^h\|_{0,\Omega}$.

2.4.2 The Convection-Diffusion Case

We now return to the convection-diffusion case, where we have, in place of (2.30),

$$\begin{aligned} \|u - u^h\|_{0,\Omega} &\leq \|u - u_0\|_{0,\Omega} + \|u_0 - u_0^h\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega} && (2.43) \\ &\leq C_1 \frac{\delta}{\sqrt{\epsilon}} + C_2 \frac{h^2}{\sqrt{\epsilon}} \|f\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega}. && (2.44) \end{aligned}$$

The $1/\sqrt{\epsilon}$ is due to the fact that the norm of the solution grows at the rate of $\sqrt{\epsilon}$, as we will verify later.

In the symmetric case, $a_{ij} = a_{ji}$, it was shown in [30] that we can write the second term for U_1^h in (2.42) in a difference form:

$$\left(G_0^h A_1^h U_0^h\right)_i = \sum_{k=1}^{N} \sum_{j=1}^{j<k} A_1^{kj} \left(U_0^j - U_0^k\right) \left(G_0^{ik} - G_0^{ij}\right); \tag{2.45}$$

$$U_1^h = G_0^h f_1^h - DG_0^h A_1^h DU_0^h. \tag{2.46}$$

D is a difference operator, details of which are not relevant to this discussion and can be found in [30]. The fact that we can write DG_0^h and DU_0^h here is crucial because this difference structure gives an additional $\mathcal{O}(h)$ in the convergence rate. With the estimates $\|G_0^h\|_{0,\Omega} \leq C/h^2$, $|f_1^h| \leq C_1\delta + C_2 h$, and $\|A_1^h\|_{0,\Omega} \leq C/h$, this leads to $\|U_1^h\| \leq C_1 + C_2/h$ in the symmetric case and hence to (2.28).

In the convection-diffusion case, we do not have the same structure. After some algebra and discrete integration by parts, we have, instead of (2.46),

$$\left(G_0^h A_1^h U_0^h\right)_i = \sum_{k=1}^{N} \sum_{j=1}^{j<k} A_1^{kj} \left(U_0^j - U_0^k\right) \left(G_0^{ik} + G_0^{ij}\right). \tag{2.47}$$

Because of the plus sign between the G_0 terms, we can write this only as

$$U_1^h = G_0^h f_1^h - G_0^h A_1^h DU_0^h, \tag{2.48}$$

without the difference operator for G_0^h term.

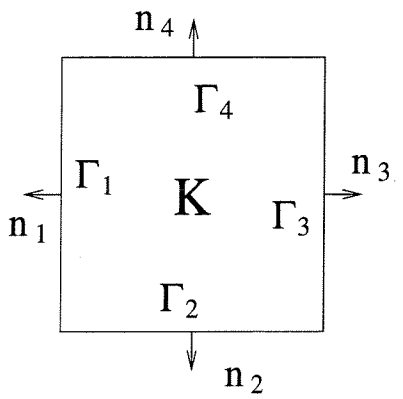

Figure 2.7: Boundary segments along an element K.

The fact that we lost $\mathcal{O}(h)$ means that the structure of (2.42) must be examined more carefully for possible additional cancellation.

2.4.3 Estimating A_1^h

We do this by carefully estimating the A_1^h term in (2.42). We denote the terms in (2.36) as follows:

$$\Lambda_1 = -\int_K \sigma^{ij} \left(\phi^k_{0,j} \theta^l_{0,i} + \phi^l_{0,j} \theta^k_{0,i} \right) dx, \tag{2.49}$$

$$\Lambda_2 = \delta \int_K a^{ij}_\delta \theta^k_{,i} \theta^l_{,j} \, dx, \tag{2.50}$$

$$\Lambda_3 = \frac{1}{\delta} \int_K \tilde{\sigma}^{ij} \phi^l_{0,j} \phi^k_{0,i} \, dx. \tag{2.51}$$

We now proceed to examine each of these terms.

i) Estimating Λ_1 :

We first move the integral to the boundary to avoid estimating $\theta^l_{,i}$ inside an element and use the fact that $\theta^l = \chi^l(\partial\phi_0/\partial x)$ on the boundary from (1.22) and that $\partial\phi_0/\partial x \sim$

$1/h$ in a cell of size h. Thus,

$$\begin{aligned}
\int_K \sigma^{ij}\phi^k_{0,j}\theta^l_{,i}\,dx &= \int_{\partial K} \sigma^{ij}\phi^k_{0,j}\theta^l n_i\,ds && (2.52)\\
&\sim \frac{1}{h}\int_{\partial K}(a^{ij}+a^{ip}\chi^j_{,y_p})\chi^l n_i\,ds && (2.53)\\
&= \frac{1}{h}\int_{\partial K} a^{ij}\chi^l n_i\,ds + \frac{1}{h}\int_{\partial K} a^{ip}\chi^j_{,y_p}\chi^l n_i\,ds. && (2.54)
\end{aligned}$$

The first term of (2.54) is $\mathcal{O}(1)$ since a^{ij} and χ^l are bounded. The second term is more difficult to estimate. Note that these terms are indexed by k, l and summed in other indices. Some terms are dropped once their sizes have been estimated, but the same indices should still be summed even though they may not be repeated. Now we group the boundary segments for the second term into two parts, with Γ_i shown in Figure 2.7. (2.54) now is the same order as

$$\begin{aligned}
&\frac{1}{h}\left(\int_{\Gamma_1+\Gamma_3} a^{1k}\chi^j_{,y_k}\chi^l n_1\,ds + \int_{\Gamma_2+\Gamma_4} a^{2k}\chi^j_{,y_k}\chi^l n_2\,ds\right)\\
&= \frac{1}{h}\left(\int_{\Gamma_1+\Gamma_3} a^{11}\chi^j_{,y_1}\chi^l n_1\,ds + \int_{\Gamma_1+\Gamma_3} a^{12}\chi^j_{,y_2}\chi^l n_1\,ds\right.\\
&\qquad \left. + \int_{\Gamma_2+\Gamma_4} a^{21}\chi^j_{,y_1}\chi^l n_2\,ds + \int_{\Gamma_2+\Gamma_4} a^{22}\chi^j_{,y_2}\chi^l n_2\,ds\right). \qquad (2.55)
\end{aligned}$$

We consider the $j = 1$ case now; the $j = 2$ case is analogous, except that the role of the second and third terms in (2.55) are reversed. The first and the last terms present little problem since the coefficient is $a^{11} = a^{22} = \epsilon$, which comes from the diffusion term and cancels out the effect of the χ term.

The second and third term appear more troublesome. χ satisfies the singularly perturbed problem (2.7) and so has layers. But we observe that $\nabla\chi$ is $\mathcal{O}(1/\sqrt{\epsilon})$, that is, the thickness of the layer in χ is $\mathcal{O}(\sqrt{\epsilon})$, as will be discussed soon. Since the stream-function ψ is a smooth function and the layers occur around the separatrices $\psi = 0$,

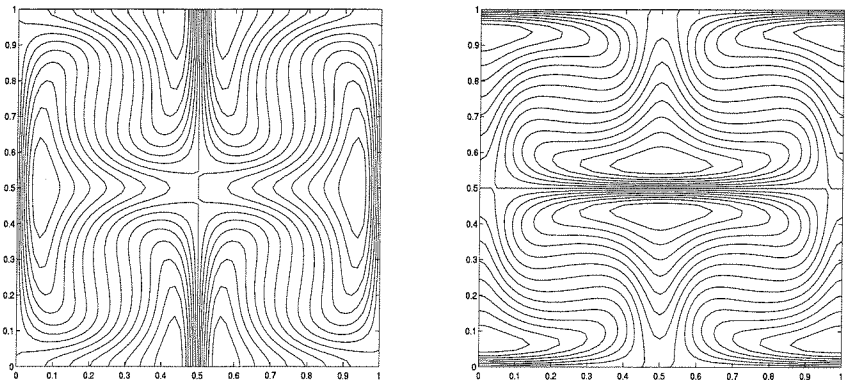

Figure 2.8: Cell problem solutions χ_1 (left) and χ_2 (right), rescaled to the unit square.

ψ is of $\mathcal{O}(\sqrt{\epsilon})$ in the layer. Thus, (2.55) can be estimated as

$$\epsilon \cdot \frac{1}{\sqrt{\epsilon}} + \int_{\Gamma_1+\Gamma_3} \psi \chi^1_{,y_2}\, ds + \int_{\Gamma_2+\Gamma_4} \psi \chi^1_{,y_1}\, ds + \epsilon \cdot \frac{1}{\sqrt{\epsilon}} \tag{2.56}$$

$$\sim \epsilon \cdot \frac{1}{\sqrt{\epsilon}} + \sqrt{\epsilon} + \sqrt{\epsilon} \cdot \frac{1}{\sqrt{\epsilon}} + \epsilon \cdot \frac{1}{\sqrt{\epsilon}} \tag{2.57}$$

$$\sim \mathcal{O}(1). \tag{2.58}$$

We plot the χ^1 and χ^2 functions in Figure 2.8 to clarify the terms in (2.56). We note, however, that the domain in the figure is over one period and not over some arbitrary element domain K of the integral. The key step in obtaining (2.58) is considering the interaction of χ and ψ. By examining where the layers occur, we obtain the $\mathcal{O}(1)$ estimate and avoid the $1/\sqrt{\epsilon}$ estimate, which is what we would get with a simpler procedure.

In (2.56) and (2.57), we have assumed that

$$\left\| \chi^l_{,j} \right\|_{L^\infty(\partial K)} \sim \frac{1}{\sqrt{\epsilon}}. \tag{2.59}$$

This is what we expect from the argument for the effective diffusivity scaling, as the gradient of χ determines the $\sqrt{\epsilon}$ scaling. In fact, we see in (2.5) that χ directly determines the σ_ϵ scaling. With the layer of width $\mathcal{O}(\sqrt{\epsilon})$ and gradient $\mathcal{O}(\sqrt{\epsilon})$, χ gives the correct scaling. We compute this numerically, and the result is shown in Table 2.2.

ϵ	σ_ϵ	ratio
0.02	0.02365	
0.01	0.01585	0.6702
0.005	0.01193	0.7527
0.0025	0.00866	0.7259

Table 2.2: $\sigma_\epsilon = \langle(\nabla\chi + \mathbf{e}) \cdot (\nabla\chi + \mathbf{e})\rangle$ scales as $\sqrt{\epsilon}$.

ϵ	$\|\chi^1_{,y_1}\|_{l^\infty}$	ratio	$\|\chi^1_{,y_2}\|_{l^\infty}$	ratio
0.02	0.7903		0.5650	
0.01	1.6755	2.1200	0.8479	1.5007
0.005	3.0552	1.8234	0.9482	1.1182
0.0025	4.7457	1.5533	1.2551	1.3236

Table 2.3: $\|\chi^1_{,y_i}\|$ has $1/\sqrt{\epsilon}$ behavior for resolved solutions over one cell.

Estimate (2.59) can also be verified directly, as shown in Table 2.3: derivatives of χ_1 along each direction have the $1/\sqrt{\epsilon}$ scaling; χ_2 gives similar results.

ii) Estimating Λ_2 :

We again transfer the integral to the boundary first.

$$\begin{aligned}
\delta \int_K a^{ij}\theta^k_{,i}\theta^l_{,j}\,dx &= \delta \int_{\partial K} a^{ij}\theta^k\theta^l_{,j}n_i\,ds && (2.60)\\
&\sim \frac{\delta}{h}\int_{\partial K} a^{ij}\chi^k\theta^l_{,j}\,ds && (2.61)\\
&\sim \frac{\delta}{h}\int_{\substack{\partial K\\ i=j}} \epsilon\theta^l_{,j}\,ds + \frac{\delta}{h}\int_{\substack{\partial K\\ i\neq j}} \psi\theta^l_{,j}\,ds && (2.62)\\
&\sim \mathcal{O}(\sqrt{\epsilon}) + \mathcal{O}(1) = \mathcal{O}(1). && (2.63)
\end{aligned}$$

In order to obtain (2.63), we use the same argument as for Λ_1, assuming that

$$\left\|\theta^l_{,j}\right\|_{L^\infty(\partial K)} \sim \frac{1}{\delta\sqrt{\epsilon}}. \tag{2.64}$$

In general, the structure of θ is very complicated, as shown in Figure 2.9, and it is hard to derive such estimates analytically. The effect of the boundary condition may travel past the immediate vicinity of the boundary. That is the reason we move the derivatives of θ from inside the element onto the boundary whenever possible. Since

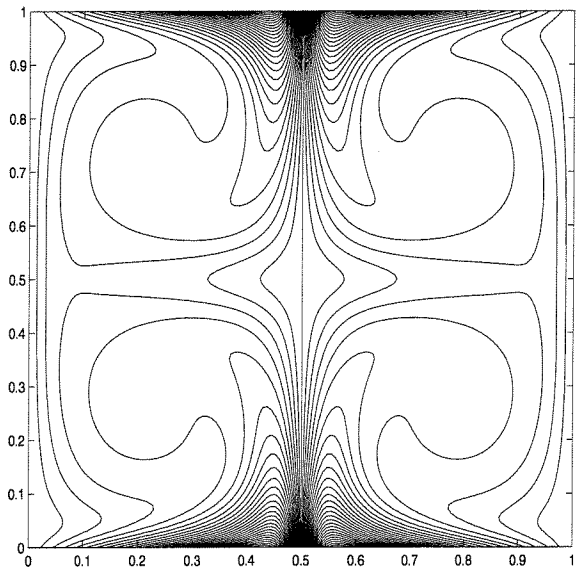

Figure 2.9: Corrector θ ($\epsilon = .001$).

ϵ	$\|\chi^1_{,y_1}\|_\infty$	ratio	$\|\chi^1_{,y_2}\|_\infty$	ratio
0.02	0.7903		0.6099	
0.01	1.6755	2.1200	1.1513	1.8877
0.005	3.0552	1.8234	1.9513	1.6949
0.0025	4.7457	1.5533	2.8846	1.4783

Table 2.4: $\|\nabla\theta\|$ has $1/\sqrt{\epsilon}$ behavior, with resolved solutions over one cell.

the estimate (2.64) is only on the boundary, the scaling is easier to understand. From (1.22), we know that θ behaves as χ on the boundary. So, the θ derivative along the boundary of the element scales as $1/\sqrt{\epsilon}$ with respect to ϵ, following (2.59). The derivative perpendicular to the boundary should not be larger than $1/\sqrt{\epsilon}$ because the layers along the boundary get weakened as they propagate inside. The numerical result in Table 2.4 appears to support this. The $1/\delta$ scaling is reasonable because the size of the periodic cell is of $\mathcal{O}(\delta)$. The $\mathcal{O}(\epsilon)$ layer within the δ-size cells gives the absolute thickness of $\mathcal{O}(\delta\sqrt{\epsilon})$.

iii) Estimating Λ_3 :

A rough estimate for Λ_3 gives

$$\frac{1}{\delta}\int_K \hat{\sigma}^{ij}\phi^l_{0,j}\phi^k_{0,i}dx = \frac{1}{\delta}\cdot\delta h\cdot\left(\frac{1}{h}\cdot\frac{1}{h}\right)\sim\frac{1}{h}. \tag{2.65}$$

ϵ	$\|G_0^h\|_{l^2}$	ratio	$\|G_0^h\|_\infty$	ratio
.04	80.6231		114.2264	
.02	144.2394	1.7891	204.3565	1.7890
.01	218.5669	1.5153	309.6592	1.5153
.005	292.8571	1.3399	414.9086	1.3399
.0025	403.3811	1.3774	571.4950	1.3774

Table 2.5: $\|G_0^h\|$ has $1/\sqrt{\epsilon}$ dependence ($n = 8, m = 256$).

However, we actually get an additional $\mathcal{O}(h)$ because we can write this term in a difference form $\Lambda_3 = D\lambda^h$ for some λ^h. Because this term does not involve χ or θ, it is exactly the same as in the elliptic case. In [30], details on the difference structures of Λ_3 are described. We can also say the same thing for f_1^h, also described in [30]. With the additional h, we conclude that $\Lambda_3 \sim \mathcal{O}(1)$.

2.4.4 Error Estimate

Putting the Λ_i estimates together, we obtain the following:

$$\|G_0^h A_1^h D U_0^h\| \leq C\left(\frac{1}{\sqrt{\epsilon}h}\right), \tag{2.66}$$

$$\|U^h - U_0^h\| \leq C_1\left(\frac{\delta}{\sqrt{\epsilon}h}\right) + C_2\left(\frac{h^2}{\sqrt{\epsilon}}\right)\|f\|_{0,\Omega}. \tag{2.67}$$

The $1/\sqrt{\epsilon}$ term comes from the definition $G_0^h \equiv \left(A_0^h\right)^{-1}$. This is the scaling we expect for G_0^h because the homogenized coefficient for A_0^h behaves as $\sqrt{\epsilon}$. We verify this numerically in Table 2.5.

Using (2.44) and the fact that $\|u\|_{0,\Omega} \sim 1/\sqrt{\epsilon}$, we finally have

$$\frac{\|u - u^h\|_{0,\Omega}}{\|u\|_{0,\Omega}} \leq C_1\left(\frac{\delta}{h}\right) + C_2 h^2 \|f\|_{0,\Omega} \tag{2.68}$$

for the relative error. Due to the discrete analysis, we get the (δ/h) term as the leading order error. This is a sharp estimate, as will be verified in the next section. We note that the relative error is independent of ϵ. If we had used the conventional finite element analysis, the first step of using Cèas's Lemma 1.2 by itself would give $1/\epsilon$, from which we would be

ϵ	δ	l^∞ norm	rate	l^2 norm	rate
.008	0.015625	3.7034		1.4633	
.004	0.015625	4.9398	1.3338	1.9515	1.3337
.002	0.015625	6.9347	1.4038	2.7391	1.4036
.001	0.015625	9.8403	1.4190	3.8858	1.4186

Table 2.6: $\|u\|$ has $1/\sqrt{\epsilon}$ behavior ($n = 4096$).

unable to recover the current estimate.

2.4.5 Numerical Results

We solve the problem numerically to verify the convergence rate given by (2.68). We use zero Dirichlet boundary condition and the forcing function

$$f = x(1 - y^2) + \sin(10(x - y)) + \cos(5(x + y)).$$

The well-resolved solution was computed on a 4096×4096 grid on the parallel computer (Intel Paragon).

In Table 2.6, we clearly see that the l^2 norm of the solution itself grows as $1/\sqrt{\epsilon}$: decreasing ϵ by 2 results in a $\sqrt{2}$ increase in the error. In order to verify the δ/h factor, we reduce δ and h at the same time. In Table 2.8, we clearly see that as δ and $h \to 0$ (with the ratio fixed at $\delta/h = 1/4$), the error stays about the same.

From (2.68), we expect the leading order error to have the $1/h$ term. We see in Table 2.7 that this is indeed the case, as halving h results in doubling of the error. In that computation, since ϵ and δ remain the same, we keep the overall resolution the same by keeping $n \times m$ constant. In all the numerical computations of this section, the results are in excellent agreement with the error estimate (2.68).

n	m	l^∞ error	ratio	l^2 error	ratio
8	256	0.0675		0.0259	
16	128	0.1605	2.3794	0.0628	2.4247
32	64	0.3358	2.0919	0.1297	2.0653
64	32	0.6134	1.8265	0.2397	1.8481

Table 2.7: $\|u - u^h\|$ has the $1/h$ term ($\epsilon = 0.008, \delta = 0.015625$).

n	δ	l^∞error	ratio	l^2error	ratio
8	0.03125	0.1112		0.04566	
16	0.015625	0.1517	0.7330	0.05954	0.7668
32	0.0078125	0.1661	0.9133	0.06457	0.9221
64	0.00390625	0.1625	1.0222	0.06365	1.0144

Table 2.8: $\|u - u^h\|$ has the δ/h term ($m = 256, \epsilon = .008$).

Choice of Trial and Test Functions

At the end of Section 2.1.2, we referred to the similarity between using the multiscale basis as the trial function (with some continuous function as the test function) and using the solution to the adjoint problem as the test function (with some continuous function as the trial function). The first method was initially motivated by homogenization to capture the small scale oscillations; the second method was simply trying to generalize the one-dimensional idea of Green's functions. However, despite the different motivations, they are very similar in their implementation.

Given these choices, it is not immediately clear what the optimal method is. We try to make some numerical comparisons in Table 2.9. We test four different cases for the cellular flow problem, each with different combinations of trial and test functions. We see that in the first case, the first three choices give almost identical results. If the trial and test functions are both linear, however, there are spurious oscillations and the method gives very large errors. In the second case (a different δ), there are some differences. Using the multiscale basis for both the trial and the test function gives about 15% less error than that given by the adjoint test function. The multiscale trial and adjoint test function combination was also tried but did not have any advantages.

U^h (trial)	V^h (test)	L^∞error	L^2error
Case I			
multiscale	multiscale	0.33347	0.11416
multiscale	linear	0.34372	0.11764
linear	adjoint	0.34344	0.11771
linear	linear	2.98428	1.08749
Case II			
multiscale	multiscale	0.20151	0.06813
multiscale	linear	0.22942	0.07766
linear	adjoint	0.23507	0.07961
linear	linear	2.98330	1.08716

Table 2.9: Error comparison for different choices of trial and test functions. Cell flow with forcing ($n = 8, m = 128, \epsilon = .008, \delta = 1/16, 1/32$).

The similar numerical experiments were performed on many different problems. Depending on the problem, different combinations give slightly smaller errors. They vary up to a factor of two on some problems, but not more. Overall, using the multiscale basis for both test and trial spaces seems to be most consistent in giving small errors. In all the cases, the error is dramatically smaller than that given by the standard bilinear functions.

2.5 Asymptotic Basis

In the multiscale FEM, the basis functions are solved numerically. So far we have assumed that we have sufficient resolution within each element to resolve all the fine structures. However, this process becomes increasingly expensive as $\epsilon \to 0$, for the same reason that the global problem became intractable in the first place. Therefore, it is natural to look for ways to reduce the work in obtaining the basis. One possibility is to look for an asymptotic expansion.

In general, this is a very difficult problem. There are many complicated internal layers along the separatrices of the streamfunction. Fortunately, we can simplify this one step by again using the asymptotic expansion:

$$\phi_\delta = \phi_0 + \delta \chi^k \left(\frac{x}{\delta}\right) \frac{\partial \phi_0}{\partial x_k} + \delta \theta^\delta. \tag{2.69}$$

This is a multiple scale expansion for the basis function, just as we had for the global solution in (1.23). Thus, we can find the asymptotic expansion for χ instead of ϕ_δ, and then use (2.69) to get an expression for ϕ_δ. In general, constructing ϕ_δ by finding χ first is not a practical option, since there may not even be a periodic structure. However, we carry out this step here to gain insights into the solution structure.

The problem now is to find the asymptotic solution to the cell problem (2.7) for χ over the torus. This is still a difficult problem with possibly very complicated layer structures. For the cellular flow, however, we can make some progress. The rest of this section is devoted to this problem.

First, we can reduce (2.7) to a "quarter-cell problem," by using certain symmetry properties. A "cell" was shown in Figure 2.3; it contains four square regions and so a quarter-cell has boundary layers along the edges and no internal layer. Looking at (2.7), when $\mathbf{e}$ is unit vector in y, χ is even in the x direction and odd in the y direction. Then by defining $\rho = \chi + x$, the analysis of (2.7) is reduced to that of [31],

$$-\epsilon\Delta\rho + \mathbf{b}\cdot\nabla\rho = 0, \tag{2.70}$$

$$\rho(x,0) = 0, \quad \rho(x,1) = 1, \quad \frac{\partial\rho}{\partial x}(0,y) = \frac{\partial\rho}{\partial x}(1,y) = 0. \tag{2.71}$$

The Boundary Layer Coordinates

To understand (2.70), we need to transform to a new coordinate system. The boundary layer structure is still too complicated to be described in the (x,y) coordinates. However, since the layers occur along the streamlines of Figure 2.3, the streamfunction $\psi(x,y)$ provides a good basis for the new coordinate system. We define $\psi(x,y)$ as one coordinate and then define a family of lines orthogonal to the streamlines [24, 78] by

$$\phi(x,y) = C\int \mathbf{b}\cdot d\mathbf{l}, \tag{2.72}$$

where $d\mathbf{l}$ is the tangent to the streamlines. This is an "angle" variable that measures the distance along a given streamline. $C = C(\psi)$ is a constant for the given level set of ψ. It

is introduced here to calibrate $\phi(x, y)$, to make sure that it takes the same range of values on each streamline. We can pick $C = 1/C_1$, where C_1 is the circulation along the given streamline. For very small ϵ, we can also choose C_1 to be circulation along the cell boundary since the layers are close to it, and still have small errors [3]. This way, C is constant for all $\psi(x, y)$ and we avoid computing it for each streamline.

With these new variables, (2.70), after some rescaling, becomes [24],

$$\frac{\partial \rho}{\partial \phi} = \epsilon \frac{\partial^2 \rho}{\partial \psi^2}, \tag{2.73}$$

with the boundary conditions

$$\rho(\phi + 4, \psi) = \rho(\phi, \psi), \tag{2.74}$$

$$\rho(\phi, 0) = 0, \quad 0 \leq \phi \leq 1, \tag{2.75}$$

$$\frac{\partial \rho(\phi, 0)}{\partial \psi} = 0, \quad 1 \leq \phi \leq 2, \tag{2.76}$$

$$\rho(\phi, 0) = 1, \quad 2 \leq \phi \leq 3, \tag{2.77}$$

$$\frac{\partial \rho(\phi, 0)}{\partial \psi} = 0, \quad 3 \leq \phi \leq 4, \tag{2.78}$$

$$\rho(\phi + 2, \psi) = -\rho(\phi, \psi). \tag{2.79}$$

Physically, this corresponds to the temperature distribution generated by periodic segments of heating, insulation, cooling, and insulation.

The solution to this problem can be obtained by the Wiener-Hopf method [3, 22]. In the interval $\phi = [0, 1]$,

$$\begin{aligned} \rho(\phi, \psi) = & -A_0 \operatorname{erfc}\left(\frac{\psi}{2\sqrt{\phi\epsilon}}\right) \\ & + \frac{1}{2\sqrt{\pi\epsilon\phi}} \int_0^\infty \left\{ \exp\left[-\frac{(\psi - \psi')^2}{4\epsilon\phi}\right] + \exp\left[-\frac{(\psi + \psi')^2}{4\epsilon\phi}\right] \right\} R(\psi')\, d\psi', \end{aligned} \tag{2.80}$$

where $R(\psi)$ is the initial distribution with $R(\psi) = \rho(0, \psi)$. We use *erf* and *erfc* to denote

the error function and the complimentary error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt, \qquad \mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt.$$

In the interval $\phi = [1, 2]$, we use $\rho(1, \psi)$ as the initial distribution and obtain

$$\rho(\phi, \psi) = \frac{1}{2\sqrt{\pi\epsilon\phi}} \int_0^\infty \left\{ \exp\left[-\frac{(\psi - \psi')^2}{4\epsilon\phi}\right] + \exp\left[-\frac{(\psi + \psi')^2}{4\epsilon\phi}\right] \right\} \rho(1, \psi')\, d\psi'. \tag{2.81}$$

Now we can get an equation for $R(\psi)$ by using (2.80), (2.81), and the fact that $\rho(2, \psi) = -R(\psi)$ from (2.79).

With the stretched variable $\xi = \psi/(2\sqrt{\epsilon})$, we then obtain the integral equation for $R(\xi)$:

$$R(\xi) = -F(\xi) + \int_0^\infty G(\xi, \xi') R(\xi')\, d\xi', \tag{2.82}$$

where

$$F(\xi) = \frac{1}{\sqrt{\pi}} \int_0^\infty \left\{ \exp\left[-(\xi - \xi')^2\right] + \exp\left[-(\xi + \xi')^2\right] \mathrm{erfc}(\xi') \right\} d\xi', \tag{2.83}$$

$$G(\xi, \xi') = \frac{1}{\sqrt{2\psi}} \left\{ \exp\left[\frac{-(\xi - \xi')^2}{2}\right] \mathrm{erf}\left[\frac{(\xi + \xi')}{\sqrt{2}}\right] + \exp\left[\frac{-(\xi + \xi')^2}{2}\right] \mathrm{erf}\left[\frac{(\xi - \xi')}{\sqrt{2}}\right] \right\}. \tag{2.84}$$

If we can find $R(\psi)$ from the integral equation (2.82), we know the solution to the problem (2.73), and hence to (2.70). This is still a difficult problem, but it was noted that the solution may be written as an expansion in temperature waves [3], and their coefficients can be computed from the analytic expression [3] given by

$$\begin{aligned}
\rho(\phi, \psi) &= \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} c_{2n-1} \exp\left(-\frac{\xi}{\delta_{2n-1}}\right) \sin\left(\frac{\xi}{\delta_{2n-1}} - (2n-1)\frac{\pi}{2}\phi - \theta_{2n-1}\right) \\
&\approx 1.0808 \exp\left(-\sqrt{\pi\xi}\right) \sin\left(\sqrt{\pi\xi} - \pi\phi/2 - 0.2949\right) \\
&\quad + 0.2536 \exp\left(-\sqrt{\pi\xi}\right) \sin\left(\sqrt{\pi\xi} - \pi\phi/2 - 0.1723\right) \\
&\quad + 0.2036 \exp\left(-\sqrt{\pi\xi}\right) \sin\left(\sqrt{\pi\xi} - \pi\phi/2 - 0.1288\right) + \cdots .
\end{aligned}$$

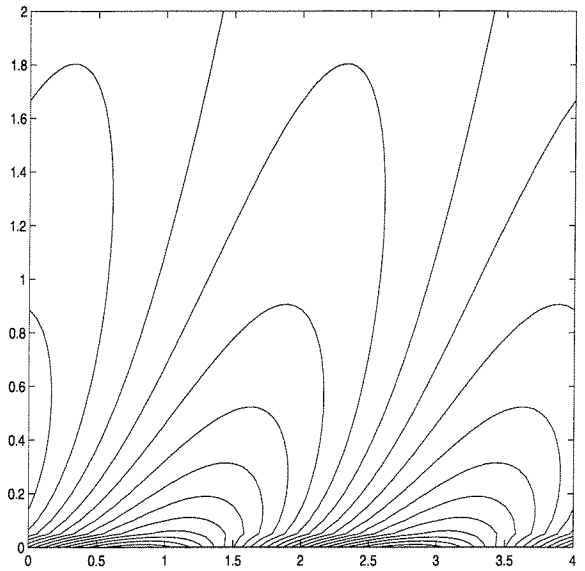

Figure 2.10: A quarter cell of χ in (ϕ, ξ) using the asymptotic solution.

This function is plotted in Figure 2.10 in (ϕ, ξ) coordinates.

This asymptotic expansion can now be converted numerically to (x, y) coordinates. In Figures 2.11 and 2.12, we have plotted the two solutions of χ using the asymptotics and the finite element method, respectively. The asymptotic expansion gives a very good approximation in this case.

By working out the asymptotic approach for the construction of the basis function in this section, we see that such an approach is possible, but only for simple problems. The multiscale basis functions in general contain complicated, genuinely two-dimensional structures. That is the reason for the success of the multiscale method, as it captures the complicated small scale features that other simpler approaches cannot. However, it means that the asymptotics becomes very difficult as well. Sometimes it can be done, as we demonstrated for the cellular flow case in this section. The boundary layer coordinates based on (2.72) may often be successful because of the generic behavior that the layers form along some streamlines. The advantage of the asymptotic approach is that the workload does not change even for very small ϵ. In general, it is difficult to carry out the asymptotics and one must resort to a numerical scheme.

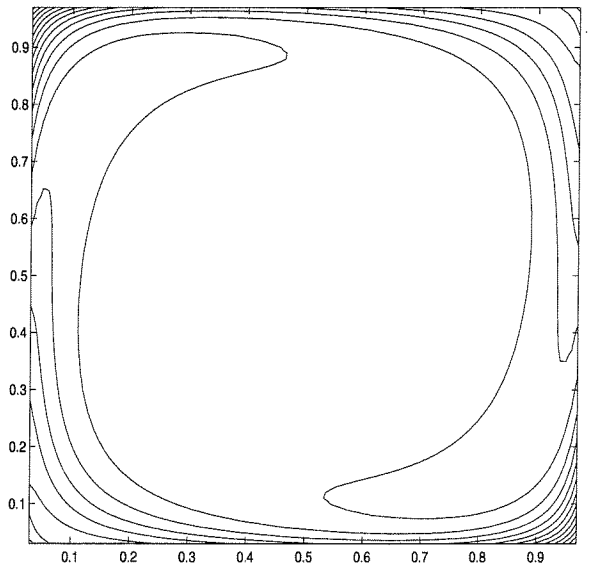

Figure 2.11: An asymptotic solution mapped to the (x, y) coordinates.

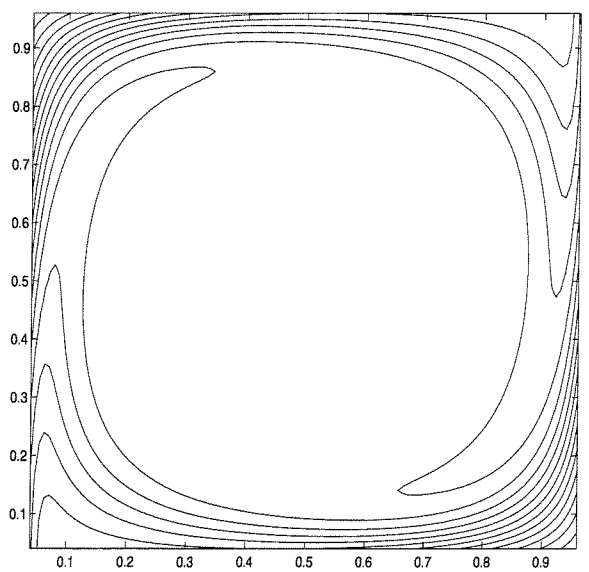

Figure 2.12: A quarter cell of χ solved numerically ($n = 128$).

2.6 Summary

We have considered in the chapter a finite element method in which the basis functions are solutions to the homogeneous equation. We were motivated by the one-dimensional results which showed that when the basis functions contain the property of the differential operator, solutions with sharp features can be computed correctly with a coarse grid. In two dimensions, we found that the multiscale finite element method works nicely when the correct behavior can be captured locally, inside each element. In particular, we have studied the cellular flow problem in detail. By resolving the fine layer structure within each element, we were able to compute the effective diffusivity scaling for the long-time, large-domain limit efficiently. Asymptotic error analysis has been carried out to show that the method converges for a fixed mesh, independent of the layer structure, if the element is larger than the small scale of the problem. This has been confirmed by numerical experiments.

Chapter 3 The Convection-Diffusion Equation with Random Coefficients

3.1 Introduction

We now turn to the more difficult case in which the characteristic length scale of the problem is large. The method used in the previous chapter does not work anymore because the basis functions that sample only a small region cannot capture the correct behavior. This nonlocality was explained in Section 1.2.3. One velocity field with such a property is the random flow constructed as a perturbation to the cellular flow. This is the velocity field we consider throughout this chapter. We first study the scaling properties, and then look for an efficient method for capturing the layers in the solution.

The properties of random velocity fields have been studied for a long time [42, 48, 74, 56, 19], particularly in relation to turbulent flows [37, 11, 32, 7]. With regard to the effective diffusivity of a passive scalar, a very interesting theoretical result was obtained previously. The velocity field is assumed to be steady, incompressible, have mean zero, and come from a stationary streamfunction $\psi(x, y)$; how this particular "random" flow is generated is described in the next section. Isichenko and Kalda [45, 47] argued that as $\epsilon \to 0$,

$$\sigma_\epsilon \sim \epsilon^{3/13}. \tag{3.1}$$

This 3/13 enhancement is a substantial one, even compared to the 1/2 in the cellular flow case. For example, when $\epsilon = .0001$, we have $\epsilon^{1/2} = .01$ and $\epsilon^{3/13} \approx .1194$. This means the overall effect of the random convection for this ϵ is more than a thousand–fold increase in the diffusivity compared to the pure diffusion and a ten–fold increase compared to the cellular flow.

This claim (3.1) is due to a scaling argument based on some results from the percola-

tion theory regarding the geometry of the streamlines of the flow, as will be described in Section 3.2.2, [46]. This scaling behavior is difficult to verify computationally, as it is an asymptotic result that is valid when the parameters involved are small. In order to approximate the infinite domain limit, which is necessary for the validity of the percolation estimates, δ must be small. δ is the parameter that rescales the domain and time in (2.1). The scaling is also in the limit of small diffusivity ϵ. In Section 3.3, we compute this problem directly on a parallel machine. We find that even for moderately small values of δ and ϵ, we can see the effect of the randomness of the flow in the enhancement of σ_ϵ.

To overcome the problem caused by the nonlocality, we need to have some knowledge regarding the structure of the layers before solving the problem. In Section 3.4, we introduce a variational principle [34], originally developed for the periodic flow and then used for random flows [31, 33]. We find that we can recover the scaling results reasonably well using this variational principle as a numerical scheme.

With the information provided by the variational principle, we attempt to design a more efficient numerical method. The idea is similar to that of the previous chapter. We would like to have the trial or test functions that contain the same behavior as the solution. In Section 3.5, we show how this idea works in one dimension for a coefficient with a turning point. We find that the one-dimensional variational principle correctly predicts the scaling of the exponential layer, and we can get a much improved solution if the exponential behavior is incorporated in the correct way. This idea is extended to two dimensions in Section 3.6. The implementation is more complicated but the insights gained from the turning point problem can be used to design an efficient method.

3.2 Random Flow Computations

3.2.1 Random Streamfunction

There are many ways of generating a random flow [4]. One common approach is to specify the spectrum in the Fourier space, as is usually done in turbulence literature [56]. Balk and

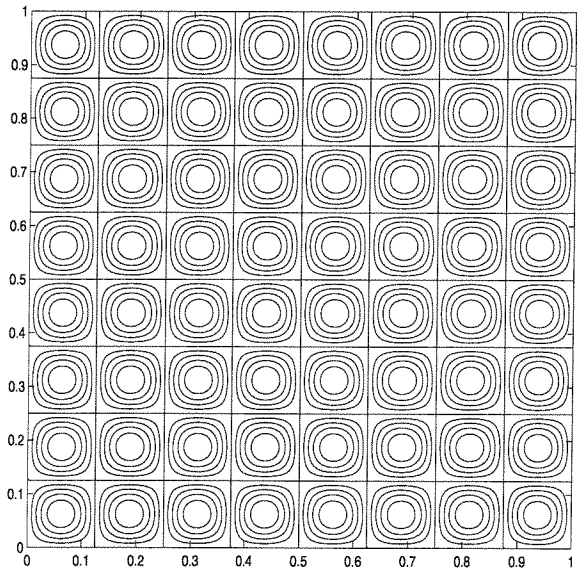

Figure 3.1: A periodic streamfunction ($\delta = .25$).

McLaughlin [11] generated it through a superposition of shear flows in random directions. Avellaneda *et al.* [8] considered a large number of vortices randomly placed. Isichenko [45] generated it by adding random perturbations to the cellular flow streamfunction. We choose the last approach here because it allows us to make use of various scaling arguments and come up with a prediction for the effect of the randomness.

For a cellular streamfunction, shown in Figure 3.1, the straight lines are the separatrices and their intersections are the hyperbolic stagnation points. Due to the nature of hyperbolic points, the cellular structure is unstable when these points are perturbed. This results in the cell boundaries reconnecting in some random manner, forming channels and islands of various sizes. Numerically, we generate a set of random numbers in the range $[-\gamma, \gamma]$ at the hyperbolic points, connect them with a smooth bicubic spline interpolation, and then add to the cellular streamfunction. A similar perturbation can be made in the Fourier space, but that approach seems to offer less control in specifying the flow. One realization is shown is Figure 3.2.

We note that in the random flow case, the cell problem (2.6) is not as useful. In the periodic case, the cell problem allows us to understand the property of the flow by looking at a single period; in the random flow, that period becomes infinitely large. However, we

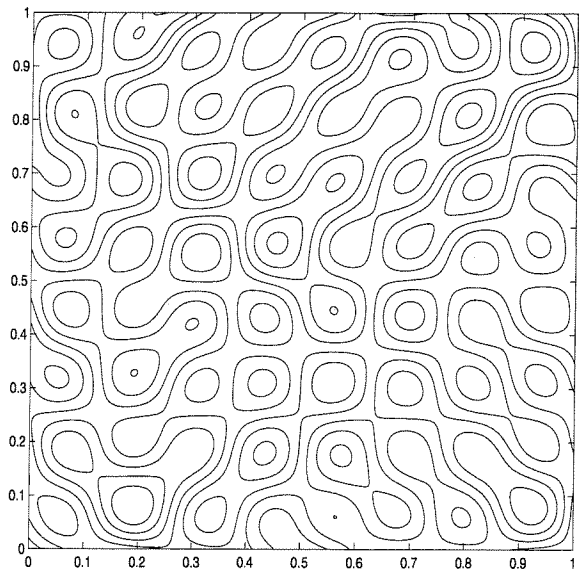

Figure 3.2: A random streamfunction ($\delta = .25, \gamma = .5$).

find it convenient to work with the cell problem, even though the cell is the whole domain, since we have the convenient expression for σ_ϵ in terms of χ, (2.5). We think of the random cell problem as the infinite volume limit of the periodic cell problem.

3.2.2 Percolation Theory and Scaling Argument

When these random reconnections occur among the cells, the streamlines form closed loops of various lengths. Fortunately, the statistical properties of the resulting configurations have been studied in connection with the "bond percolation clusters," and their behavior is well-understood. The idea is to map the continuum percolation to a lattice and study how the cells are connected to their neighbors through a "bond" when the probability of each bond occurring is given. Details are given in [33] and the references therein. We simply cite the main results here.

We define the level set h of some streamfunction $\psi(x, y)$ to be $\{(x, y) \,|\, \psi(x, y) = h\}$. This level set is a complicated curve with some perimeter. From the percolation theory, the

typical "diameter" $\xi(h)$ of the level set h near $\psi = 0$ scales as

$$\xi \sim h^{-\nu}, \quad \nu = 4/3. \tag{3.2}$$

That is, the diameter grows larger as h approaches 0 and it is infinite at the critical set $h = 0$. The exponent ν is a measure of the correlation length of the cluster. The complicated perimeter of this cluster actually forms a fractal curve, with the fractal dimension of $d_h = 1 + 1/\nu = 7/4$. These values are hard to prove rigorously, but many numerical experiments support the predictions. A Monte-Carlo simulation of a large perimeter clusters in [83], for example, yields $d_h = 1.751 \pm 0.002$.

Putting (3.2) and the fractal dimension together, the typical length of the curve near $h = 0$ is

$$l(h) \sim h^{-\nu d_h}. \tag{3.3}$$

When $\nabla\psi \sim 1$, we also know [33] that the width of the contour lines with diameter of order ξ is

$$w(\xi) \sim \xi^{-1/\nu}. \tag{3.4}$$

Intuitively, these estimates mean that near the critical level $\psi(x, y) = 0$, there emerges a very long loop. This loop then gives the dominant contribution to the overall diffusivity because once a particle diffuses onto this layer, it gets transported very far before it has a chance to diffuse out. The longest zero streamline of one realization is shown in Figure 3.3. We see that as $\delta \to 0$, it will have a fractal structure as predicted.

With these estimates, we can now make a scaling argument for σ_ϵ. As in the periodic case, the diffusive time scale across the layer of width w should equal the convective time scale over the cell of size l:

$$\frac{w^2(h)}{\epsilon} \sim \frac{l(h)}{u_0}, \quad (h > 0). \tag{3.5}$$

The magnitude of velocity u_0 here is $\mathcal{O}(1)$ and the size of the cell is also $\mathcal{O}(1)$. Unlike the

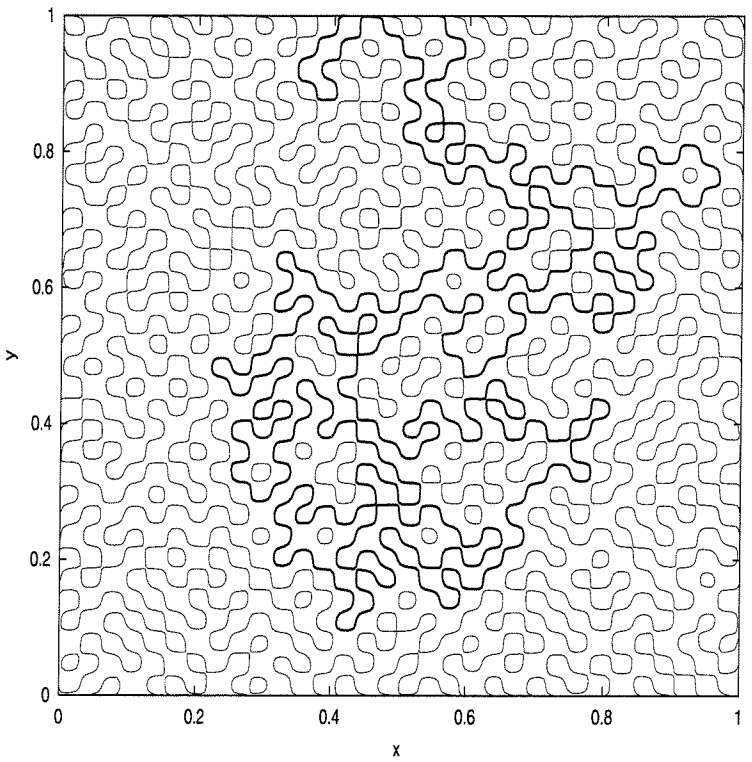

Figure 3.3: Zero level set contour lines ($\delta = .0625$, $\gamma = .5$).

periodic case, the width w and the length l of the layer now depends on the level set h. The $w(h)$ can be obtained by combining (3.2) and (3.4). This results in $w(h) \sim h$. Now substituting these in (3.5), we get

$$\frac{h^2}{\epsilon} \sim h^{-\nu d_h},$$

or,

$$h_\epsilon \sim \epsilon^{\frac{1}{2+\nu d_h}} = \epsilon^{3/13}.$$

We use h_ϵ to denote this particular level set that contains the long fractal boundary layer. From this h_ϵ, we can compute the diffusive flux using (2.5), [31].

$$\begin{aligned}
\sigma_\epsilon &= \epsilon \left\langle (\nabla\chi + \mathbf{e}) \cdot (\nabla\chi + \mathbf{e}) \right\rangle && (3.6)\\
&\sim \epsilon \left(\frac{1}{h_\epsilon}\right)^2 l(h_\epsilon)\, w(h_\epsilon) && (3.7)\\
&\sim \epsilon \left(\frac{1}{\epsilon^{3/13}}\right)^2 \epsilon^{-7/13}\, \epsilon^{3/13} && (3.8)\\
&\sim \epsilon^{3/13}. && (3.9)
\end{aligned}$$

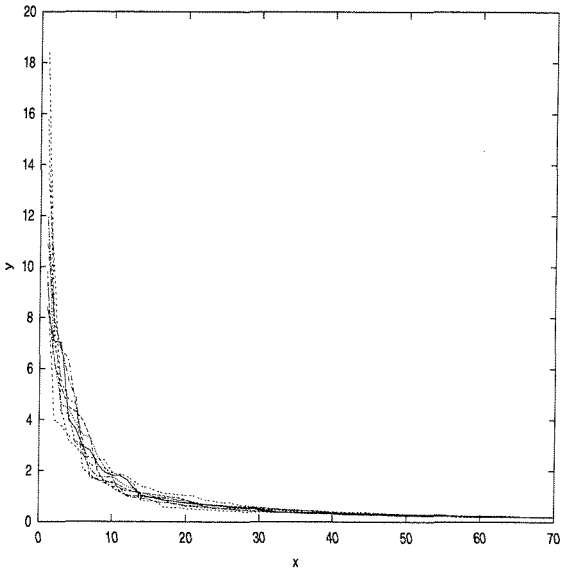

Figure 3.4: Distribution of the contour lengths for 5 realizations. y-axis is $l(h_\epsilon)$ and x-axis is the frequency of occurrence.

The main contribution (3.7) by the long loop is computed simply by multiplying the gradient in the layer to its area $l(h_\epsilon)w(h_\epsilon)$. We see that the width of the layer h_ϵ is translated into $\nabla\chi \sim 1/h_\epsilon$ within the layer and then to the $\sigma_\epsilon \sim 3/13$ scaling.

In Figure 3.4, each curve represents a different realization of the random flow. y-axis is the length of the loops at the level set $h = 0$ and x-axis is the frequency at which the curve of given length appears. The figure shows that there are few loops with very long perimeter and the number increases as the length gets shorter. In the limit of an infinite domain, there would be a single loop of infinite size.

3.3 Large-Scale Computations

The scaling result (3.1) is a theoretical one. There has been little numerical evidence to support this claim. The reason for the lack of computational work is that it is difficult to have enough resolution to verify an asymptotic result that is in the limit of small viscosity ($\epsilon \to 0$) and infinite domain ($\delta \to 0$). More "cells" are squeezed into a given computational resolution as $\delta \to 0$ and the layers get thiner as $\epsilon \to 0$, making it difficult to achieve sufficient

resolution.

One previous numerical result for the σ_ϵ scaling is contained in [48]. In that work, a Monte-Carlo method was employed: at each time step Δt, particles are advected according to $d\mathbf{x}/dt = \mathbf{b}(\mathbf{x})$ and displaced by a distance $\sqrt{4\epsilon\Delta t}$ in a random direction. Then the distances traveled by the particles are averaged. The streamfunction used to generate the random velocity field $\mathbf{b}$ was

$$\psi_N(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sin(\mathbf{k}_i \mathbf{r}), \tag{3.10}$$

where the direction of $\mathbf{k}_i$ is chosen randomly. The $N^{-1/2}$ factor is introduced to fix the average amplitude. The computation was performed with 256 particles for different number of directions superimposed. The scaling exponent they compute is β of the form $\sigma_\epsilon \approx \epsilon(Pe)^\beta$. Since the Peclet number Pe has the $1/\epsilon$ factor, the value of $3/13$ scaling we have mentioned is equivalent to $\beta = 10/13 \approx 0.7692$. The estimates of β obtained by the Monte-Carlo approach are $0.72 \pm 0.03\,(N = 12)$, $0.81 \pm 0.03\,(N = 18)$, $0.69 \pm 0.02\,(N = 25)$ and $0.81 \pm 0.03\,(N = 30)$. These results were inconclusive, as the computation was sensitive to the parameters.

This type of Monte-Carlo approach often is satisfactory if one is interested only in the scaling result. However, our goal is not just to obtain the scaling, but also to compute the correct solution itself. The scaling computation is simply an interesting application.

Our approach in this section is a direct simulation, resolving the details with a very fine mesh, to see the scaling behavior. We do this by solving for the steady-state solution $\rho(x, y)$ given by

$$-\epsilon\Delta\rho + \mathbf{b} \cdot \nabla\rho = 0, \tag{3.11}$$

$$\rho(x,0) = 0, \quad \rho(x,1) = 1, \quad \frac{\partial\rho}{\partial x}(0,y) = \frac{\partial\rho}{\partial x}(1,y) = 0, \tag{3.12}$$

just like (2.70) and then use

$$\sigma_\epsilon \approx \epsilon \iint |\nabla\rho|^2 \, dx\, dy. \tag{3.13}$$

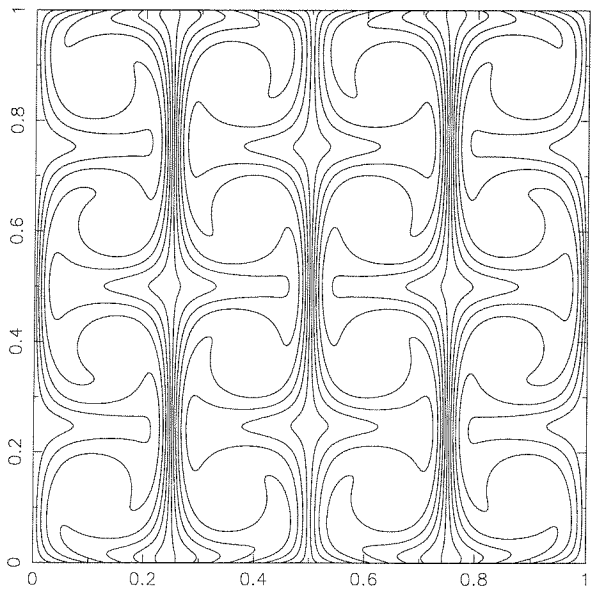

Figure 3.5: A cell flow computation ($\epsilon = .001, \delta = .5, \gamma = 0$).

This is equivalent to the cell problem (2.5) for $\chi = \rho - x$. (3.13) is an approximation because the problem is no longer periodic. In Figure 3.5, we plot ρ for the periodic cellular flow. In Figure 3.6, we see the solution after random perturbations have been added to the flow. The computations for obtaining σ_ϵ are done with smaller δ, but here we show the $\delta = .5$ case, so that the details can be observed. The size of the perturbation γ was .5. In Figure 3.7, we plot the solution for the $\delta = .125, \epsilon = .001$ case.

We could also solve the unsteady problem with a regularized delta function as an initial condition and compute σ_ϵ using (2.24), as we did for the cellular flow in Section 2.3.2. However, this presents some complications in the numerical problem. We find that in order for any structures related to the topography of the streamlines to appear, we need to run the equation for a long period of time. However, because of the increase in the effective diffusivity, we have the fluid reaching the boundary more quickly.

3.3.1 Numerical Results

We solve the steady problem (3.11). We use the Intel Paragon parallel computer as in Section 2.3.2. It has maximum of 512 nodes, each with 32 MB of memory. We are able to

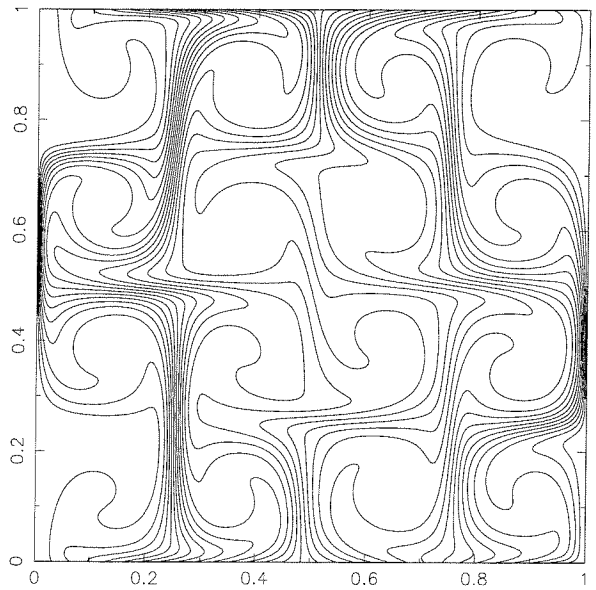

Figure 3.6: A random flow computation ($\epsilon = .001, \delta = .5, \gamma = .5$).

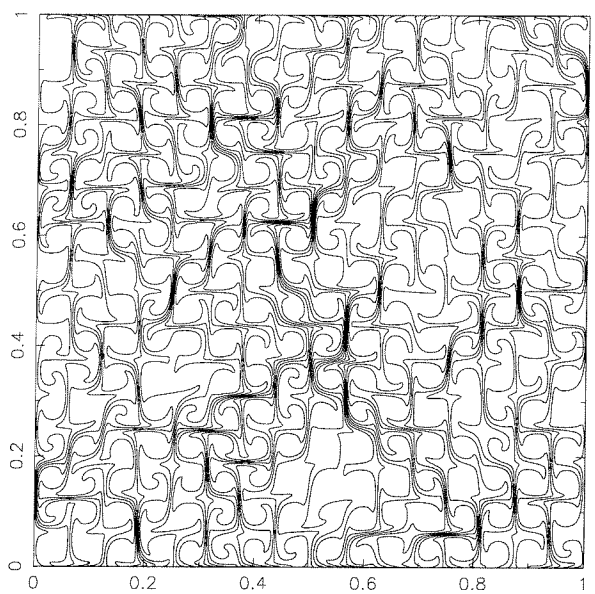

Figure 3.7: A random flow computation ($\epsilon = .001, \delta = .125, \gamma = .5$).

ϵ	δ	σ_ϵ	rate
0.00025	0.25	0.00557396	0.834036
0.00025	0.125	0.00518526	0.858599
0.00025	0.0625	0.00512630	0.850075

Table 3.1: Estimating the scaling exponent as $\delta \to 0$; we compute the rate as $\epsilon = .0005 \to .00025$ for each δ ($\gamma = .5, n = 2048$).

ϵ	δ	σ	rate
0.016	.25	0.02095145	
0.008	.25	0.01530900	0.73068
0.004	.25	0.01222705	0.79868
0.002	.25	0.00975801	0.79806
0.001	.25	0.00800031	0.81987
0.0005	.25	0.00668312	0.83535
0.00025	.25	0.00557396	0.83403

Table 3.2: Estimating the scaling exponent as $\epsilon \to 0$; rate of change as for σ_ϵ as $\epsilon \to \epsilon/2$ for a fixed δ ($\gamma = .5, n = 2048$).

compute on a grid as large as 4096×4096 on this machine. The multigrid solver [28] has been parallelized and it works fairly well for the random flow, as long as perturbation from the cellular flow is not too large.

If the correct scaling exponent is indeed $3/13$, we expect that when $\epsilon \to \epsilon/2$, σ_ϵ is reduced by

$$\left(\frac{1}{2}\right)^{3/13} \approx 0.85218. \tag{3.14}$$

In the cellular flow case, this number was $1/\sqrt{2} \approx .707107$.

In Table 3.1, we see that with small ϵ, the rate of change is very close to the predicted number as $\delta \to 0$. In theory, this number should approach the number in (3.14) as δ is reduced. It is not conclusive with three data points, but we see that it appears to support the predicted scaling. Since the scaling result is an asymptotic one in the limit of both $\epsilon \to 0$ and $\delta \to 0$, we record the rate as $\epsilon \to 0$ with δ fixed in Table 3.2. The rate is not monotone, but even with $\delta = .25$, it is increasing slowly in the right direction, to a number close to the predicted one.

3.4 Variational Principle

Variational principles have long been used in material science and other fields in order to obtain bounds on effective properties of nonhomogeneous media; see, e.g., [64]. Naturally, such approach has also been used for the effective diffusivity problem. Avellaneda and Majda in [7, 6] derive an integral representation and used it to find upper bounds. In [31], Fannjiang and Papanicolaou find another form of the variational principle and its dual for the effective diffusivity of the convection-diffusion equation. Their variational principle allows one to attain rigorous estimates of the effective diffusivity for many flows when the test functions are suitably chosen. While these variational principles have been useful for theoretical purposes, we show how they can be used as a numerical tool. In the first section, we briefly review the procedure in [31] for obtaining the variational principle. Then we use it in the next section to estimate the scaling exponent.

3.4.1 Derivation

The main difficulty in obtaining a variational principle for the convection-diffusion equation comes from the convection term. Because of the skew-symmetric part $\mathbf{H}$ in the coefficient $(\epsilon\mathbf{I} + \mathbf{H})$, we cannot easily formulate a functional that has σ_ϵ as the minimum. In order to find this functional, the problem has to be *symmetrized* in some way.

The solution to the equation has a symmetric part and an anti-symmetric part. It is then natural to separate the two parts of the solution by adding and subtracting to it the solution of the adjoint problem. This is analogous to finding the symmetric and anti-symmetric part for a given matrix.

Recall that the space $\mathcal{H}$ of square integrable vector functions can be decomposed as

$$\mathcal{H} = \mathcal{H}_g \oplus \mathcal{H}_c \oplus \mathcal{H}_0, \tag{3.15}$$

where $\mathcal{H}_g$ is the subspace of gradient fields, $\mathcal{H}_c$ the subspace of divergence-free fields, and $\mathcal{H}_0$ the subspace of constants.

For a given function $\mathbf{f}(\mathbf{x})$ in $\mathcal{H}$, we have the Fourier representation

$$\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\hat{\mathbf{f}}(\mathbf{k}), \tag{3.16}$$

with Fourier coefficients $\hat{\mathbf{f}}(\mathbf{k})$. We define Γ_g to be the orthogonal projection operator onto $\mathcal{H}_g$, given by

$$\Gamma_g\mathbf{f} = \nabla\Delta^{-1}\nabla\mathbf{f} = \sum_{\mathbf{k}\neq 0} \frac{\mathbf{k}(\mathbf{k}\cdot\hat{\mathbf{f}}(\mathbf{k}))}{|\mathbf{k}|^2} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{3.17}$$

We can understand this definition for Γ_g by the following. We can write any vector $\mathbf{f}$ as $\mathbf{f} = \mathbf{c}+\nabla\times\mathbf{f_1}+\nabla f_2$ according to (3.15), for some constant $\mathbf{c}$, a vector field $\mathbf{f_1}$, and a function f_2. We see that if we apply $\nabla\Delta^{-1}\nabla$ to $\mathbf{f}$, all terms will disappear except the gradient field ∇f_2.

With this Γ_g operator, it is possible to rewrite the cell problem (2.6) in an integral form [31], as

$$\mathbf{E}^+ = \mathbf{e} - \Gamma_g\mathbf{H}\mathbf{E}^+, \tag{3.18}$$

where we have set $\epsilon = 1$ momentarily for simplicity and defined $\mathbf{E}^+ \equiv \nabla\chi + \mathbf{e}$, the solution of the cell problem (2.6). Now we define $\mathbf{E}^-$ to satisfy

$$\mathbf{E}^- = \mathbf{e} + \Gamma_g\mathbf{H}\mathbf{E}^-. \tag{3.19}$$

Now we can separate the symmetric and anti-symmetric parts as we wanted, by setting

$$\mathbf{A} = \frac{1}{2}\left(\mathbf{E}^+ + \mathbf{E}^-\right), \quad \mathbf{B} = \frac{1}{2}\left(\mathbf{E}^+ - \mathbf{E}^-\right). \tag{3.20}$$

After several steps [31], we find that we can write (2.5) as

$$\sigma(\mathbf{e}) = \epsilon\left\langle\left(I - \frac{1}{\epsilon^2}\mathbf{H}\Gamma_g\mathbf{H}\right)\mathbf{A}\cdot\mathbf{A}\right\rangle. \tag{3.21}$$

Now we have a complicated *nonlocal* operator $I - (1/\epsilon^2)\mathbf{H}\Gamma_g\mathbf{H}$ instead of a matrix. The advantage, however, is that the operator is self-adjoint and positive definite. Thus, we have symmetrized the problem, at the cost of introducing the nonlocal operator Γ_g.

With some simplifications on (3.21), we arrive at the form of two-dimensional variational principle we can use:

$$\sigma_\epsilon(\mathbf{e}) = \inf_{\langle \nabla f \rangle = \mathbf{e}} \left(\epsilon \langle \nabla f \cdot \nabla f \rangle + \frac{1}{\epsilon} \langle \Gamma_g \mathbf{H} \nabla f \cdot \Gamma_g \mathbf{H} \nabla f \rangle \right). \tag{3.22}$$

3.4.2 Variational Principle as a Numerical Method

We notice that in the scaling argument of Section 3.2.2 for determining σ_ϵ (3.6), the major contribution to the integral comes from the region of large $\nabla\chi$. The integral is basically the value of $|\nabla\chi|^2$ in the boundary layer multiplied by the area of the layer $w(h_\epsilon)\, l(h_\epsilon)$.

Given this information, it is reasonable to try test functions of the form

$$f(x,y) = \exp\left(-\frac{|\psi(x,y) - c|}{\epsilon^\alpha} \right) \tag{3.23}$$

for approximating the behavior of χ. α is the unknown exponent we wish to find and c is the unknown level set at which the longest streamline occurs. We expect this c to be near 0 if the other parameters are such that the percolation results hold. Since $f(x,y)$ has a layer of thickness proportional to ϵ^α and $\nabla f \sim 1/\epsilon^\alpha$, we expect the α of (3.23) that achieves the minimum of the variational integral to be close to the true value of the exponent.

Thus, in our new numerical method, we vary α and c until the minimum of the integral is reached. It involves starting with some initial α and c in the test function in (3.23), evaluating the variational integral (3.22), and then repeating this process with suitable choices of α and c until the minimum is found.

For this purpose, we can rewrite (3.22) in a form more amenable to computation. We would like to avoid working with Γ_g operator. We can do this by introducing a new function g

that satisfies

$$\Gamma_g \mathbf{H} \nabla f = \epsilon \nabla g. \tag{3.24}$$

Then, with $\Gamma_g \equiv \nabla \Delta^{-1} \nabla$ and $\nabla \cdot \mathbf{H} = -\mathbf{b}$, (3.24) simplifies to

$$-\epsilon \Delta g = \mathbf{b} \cdot \nabla f. \tag{3.25}$$

This is a Poisson problem for g with forcing $-(\mathbf{b} \cdot \nabla f)/\epsilon$. Now the variational principle can be written as

$$\sigma_\epsilon(\mathbf{e}) = \inf_{\langle \nabla f \rangle = \mathbf{e}} \left(\epsilon \langle \nabla f \cdot \nabla f \rangle + \epsilon \langle \nabla g \cdot \nabla g \rangle \right), \tag{3.26}$$

where f and g are related by (3.25).

Thus, given some f, we iterate through the process of solving the Poisson problem (3.25) for g and then evaluating (3.26). The Poission equation can be solved fast and the evaluation of the integral only involves some differentiations and summations.

We note that the true minimizer of this integral is

$$f_i = \frac{1}{2}(\chi_i^+ + \chi_i^-) + x_i, \tag{3.27}$$

where χ_i^+, χ_i^- are the solutions of

$$-\epsilon \Delta \chi_i^+ + \mathbf{b} \cdot \nabla \chi_i^+ = -\mathbf{b} \cdot \mathbf{e}_i,$$
$$-\epsilon \Delta \chi_i^- - \mathbf{b} \cdot \nabla \chi_i^- = \mathbf{b} \cdot \mathbf{e}_i,$$

with periodic boundary conditions. χ_i^+ is the solution of the cell problem (2.7) introduced earlier and χ_i^- is the solution of the same problem but with the flow going in the opposite direction. The subscript i can be 1 or 2, corresponding to the two possible directions of the unit vector $\mathbf{e}$. The relationship between f_i and g_i is analogous to that of $\mathbf{A}$ and $\mathbf{B}$ in (3.20). With f defined as (3.27), we can find that $g_i = -1/2(\chi_i^+ - \chi_i^-)$. With some algebra, we can directly verify that $\epsilon \langle \nabla f \cdot \nabla f \rangle + \epsilon \langle \nabla g \cdot \nabla g \rangle = \sigma_\epsilon(\mathbf{e})$, using the fact that

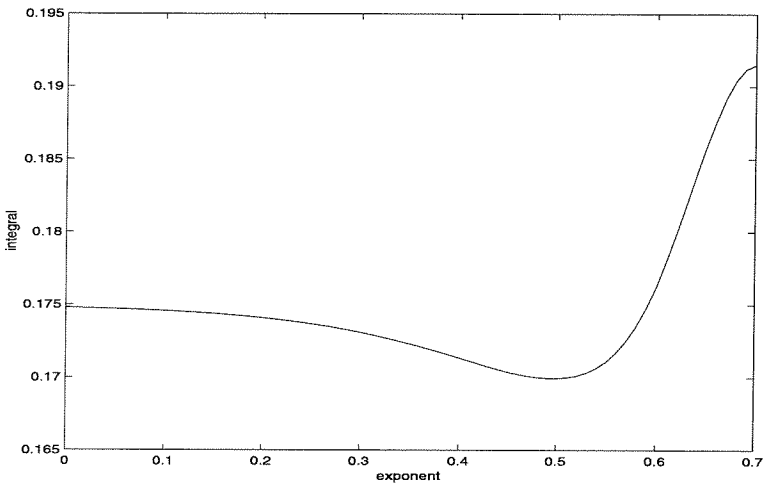

Figure 3.8: Variational integral for the cellular flow; minimum is achieved at $\alpha = .5$ as expected.

$$\sigma_\epsilon = \epsilon \left\langle (\nabla \chi_i^+ + \mathbf{e}_i) \cdot (\nabla \chi_i^+ + \mathbf{e}_i) \right\rangle = \epsilon \left\langle (\nabla \chi_i^- + \mathbf{e}_i) \cdot (\nabla \chi_i^- + \mathbf{e}_i) \right\rangle.$$

While the test function (3.23) we use is not the true minimizer (3.27), we are mimicking the essential behavior with a simple artificial function.

3.4.3 Periodic Case

We test the method for the cellular flow problem first. We know, both from analytical and numerical result, that α should be $1/2$ in this case. We plot the value of the integral as a function of α in Figure 3.8. We see that it gives a correct estimate for α. This gives some confidence that our heuristic approach is valid. We now test it for the random flow.

3.4.4 Random Case

The random case is more sensitive to the parameters and therefore requires more attention. In the periodic case, for example, the same estimate for α is achieved for different values of δ, unlike the random case. In the small diffusivity, infinite domain limit, the exponent should be approaching $3/13 \approx 0.2307$. We do not expect to obtain this number exactly in a finite-domain computation with limited resolution. But we see in Table 3.3 that as δ

ϵ	δ	exponent
0.0001	1	.34
0.0001	.5	.26
0.0001	0.25	.23
0.0001	0.125	.22

Table 3.3: Estimate of the scaling for the random flow with the variational principle; at the exponent listed, the variational integral achieves the minimum.

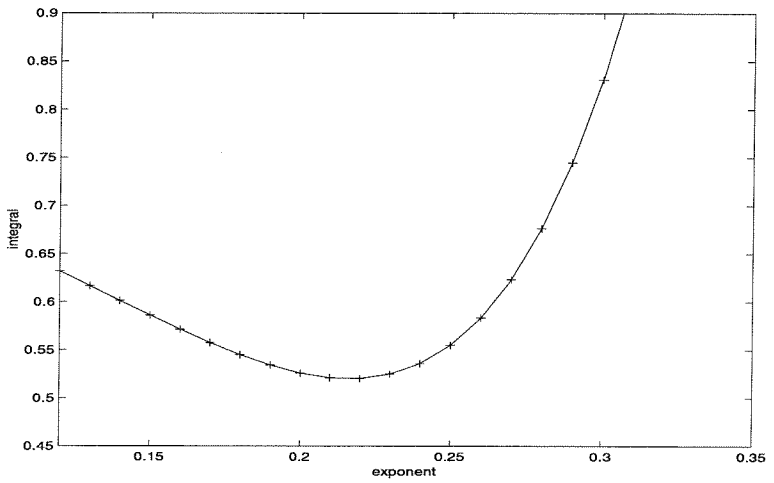

Figure 3.9: Variational integral for the random velocity field as a function of the exponent α ($\epsilon = .0001$, $\delta = .125$).

decreases, the α at which the functional is minimized is decreased and is close to the limit. At $\delta = .125$, it goes below the expected number. We see that even for a large δ, there is a substantial change from the cellular case. A typical shape of the functional is shown in Figure 3.9. In the test function (3.23), we also need to find the value of c. This is the level set of the streamfunction where the layer of interest occurs. We expect this number to be close to zero in our computations. In Figure 3.10, we show the value of the integral as a function of c. We indeed see that the energy is minimized at $c \approx 0$.

The idea presented in this section can be applied more generally. Given a problem and its variational principle, if the dominant contribution to the integral comes from a layer structure of the solution, we can use this idea to find out the thickness and the location of the layer. Then we can use this information to design a more efficient method to solve the original problem. This is illustrated in the next two sections.

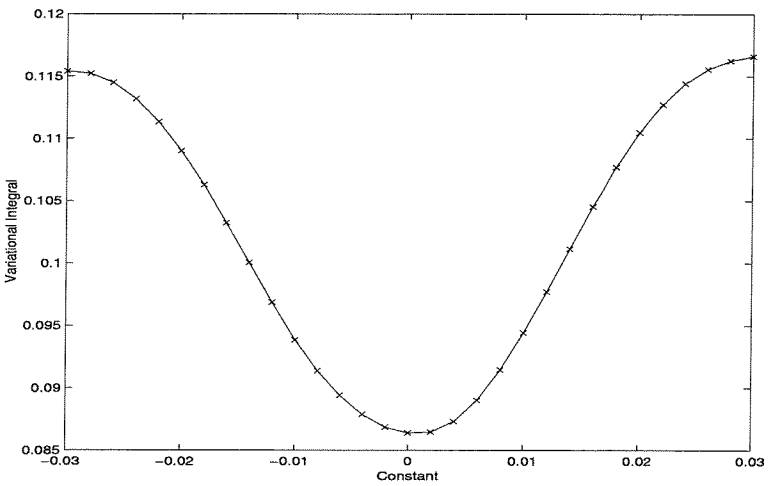

Figure 3.10: Variational integral for the random velocity field as a function of the constant c ($\alpha = .23$, $\epsilon = .0001$, $\delta = .25$).

3.5 Exponential Basis Functions: One-Dimensional Problem

Many authors have studied the one-dimensional convection-diffusion problem with variations on the exponential basis idea. We noted in Section 2.1.1 that the test function space should be composed of exponentials that solve the local adjoint problem in order to minimize the error for a given mesh. However, all these methods have assumed a positive coefficient, $a(x) \geq a_0 > 0$, [73, 68, 63, 81].

In this section, we study how basis functions should be constructed for a more general case when the coefficient has a turning point. Understanding the turning point case in one dimension is essential for the two-dimensional problems studied in the next section, since a similar phenomenon occurs all along the layers.

We consider the simple case

$$-\epsilon u'' + a(x)u' = 0, \qquad a(x) = \begin{cases} 1, & 0 < x < 1/2 \\ 0, & x = 1/2 \\ -1, & 1/2 < x < 1 \end{cases} \tag{3.28}$$

with boundary conditions $u(0) = 0$ and $u(1) = 1$. We can find the exact solution by solving

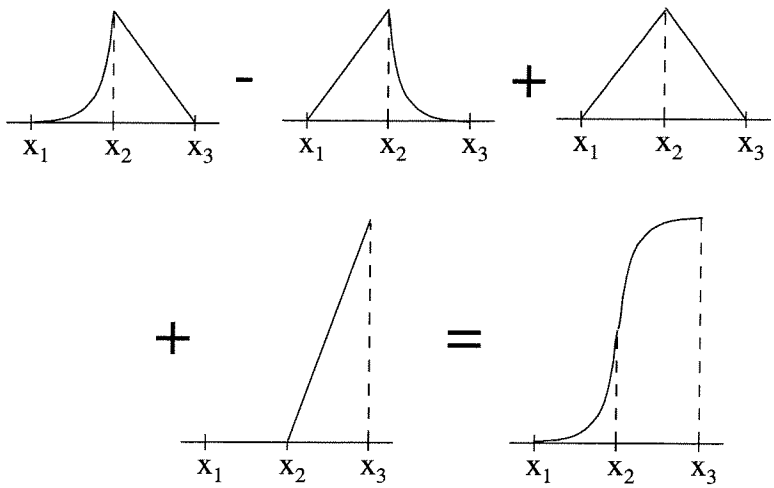

Figure 3.11: A proper combination of basis functions for an internal layer: two bases with the exponential behavior on different sides, a hat function over the same support, and another hat function in the next support are needed to achieve the shape of the layer shown.

the equation in each region and imposing continuity of the solution and its derivative at $x = 1/2$. The solution is

$$u(x) = \begin{cases} \frac{e^{x/\epsilon}-1}{2\left(e^{1/2\epsilon}-1\right)}, & 0 < x \leq 1/2 \\ 1 - \frac{e^{(1-x)/\epsilon}-1}{2\left(e^{1/2\epsilon}-1\right)}, & 1/2 \leq x < 1. \end{cases} \tag{3.29}$$

There is a sharp internal layer at $x = 1/2$. Intuitively, the "fluid" on the left of $x = 1/2$ is being pushed to the right while the opposite is true on the right of $x = 1/2$.

For this problem, simply replacing the linear basis functions with the local solutions of the homogeneous equation does not work. We still need to solve the equation locally but we need to incorporate this information in a different way. It turns out that we must separate the exponential behavior on each side of the turning point into two different bases and include all the linear functions as well. This is illustrated in Figure 3.11. We see that at the turning point x_2, there are three basis functions over the same support: one linear, one with the exponential layer on the left, and the last with the exponential layer on the right. The problem is well-posed as long as the three bases remain linearly independent. This is true if the diffusion is weak and the grid is relatively coarse.

Once we have the necessary basis functions, we need to reformulate the problem accounting

for the overlapping bases. Assuming that there are two types of bases, ϕ_i and ψ_i, that are linear and exponential respectively, we write

$$u = \sum_{i=1}^{n} a_i \phi_i + \sum_{i=1}^{2n} b_i \psi_i. \tag{3.30}$$

When we know where the layer will occur, we only need to include ψ_i where they are needed.

One way to solve this is as a minimization for the variational principle for (3.28), given by

$$I(u(x)) = \min_{\substack{u \in [0,1] \\ u(0)=0 \\ u(1)=0}} \frac{1}{2} \int_0^1 e^{-b\epsilon} \left((u_x)^2 - \frac{2uf}{\epsilon} \right) dx, \tag{3.31}$$

where $b(x) \equiv \int^x a(s)\, ds$. The validity of (3.31) can be seen easily by setting $\delta I = 0$ and finding the corresponding Euler equation.

We now substitute (3.30) in (3.31) and impose the boundary conditions to get a minimization problem. This can be solved, for example, by the conjugate gradient method. We can find explicit expressions for the terms in ∇I and use them to move down the gradient. Numerical experiments show that for moderate size n, correct weights a_i and b_i are found in a short amount of time.

While using the one-dimensional variational problem (3.31) was motivated by the previous section, it is not the best way to find a_i and b_i. The best way of solving this problem is through the finite element formulation, which can also be thought of as a variational problem. We find that we are able to handle many unknowns more easily that way.

The finite element method now needs to be modified for the extra unknowns b_i. When ψ_i are present, the local stiffness matrix is not 2×2 anymore. If we were to keep all $2n$ extra bases in (3.30), the global stiffness matrix changes from $n \times n$ to $3n \times 3n$. In general, we include the extra bases only near the possible layer locations. Therefore, the matrix problem increases only by a small number of rows and columns. While the new matrix does not have a banded structure anymore, it can be solved efficiently, as will be described in Section 3.6.3.

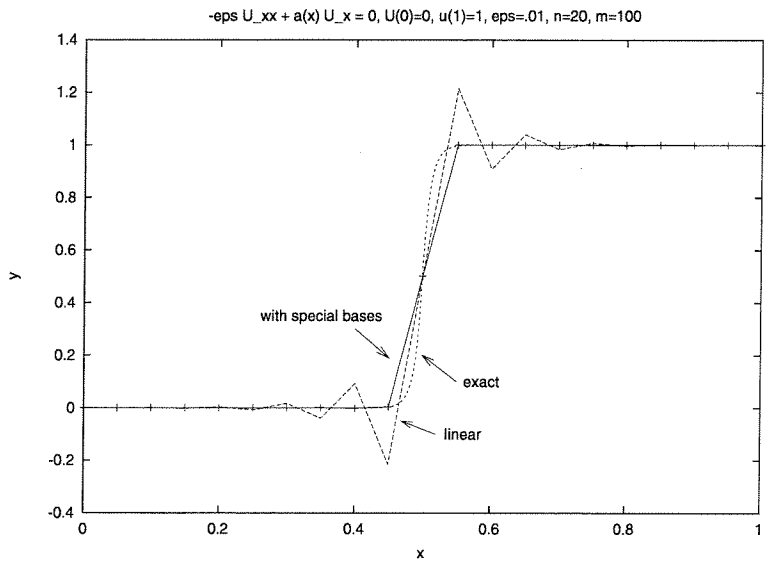

Figure 3.12: 1-D internal layer problem with and without extra bases: using linear bases only results in oscillations; combining with exponential bases in a proper combination results in exact nodal values for the coefficient in (3.28).

In Figure 3.12, we see that using the linear basis alone gives oscillations when the mesh is not fine enough. With the extra basis, not only are the oscillations suppressed, the nodal values are exact for this turning point problem with piecewise constant coefficient (3.28). As $\epsilon \to 0$, the oscillations get larger with only the linear bases, while the new approach performs the same regardless of ϵ.

3.6 Exponential Basis Functions: Two-Dimensional Problem

In two dimensions, we would also like to design basis functions that capture the exponential behavior. The difficulty, as mentioned in Section 3.1, is that solving the equation locally does not give the correct behavior for the velocity fields we consider in this chapter. This is the reason we introduced the variational principle in Section 3.4. In this section, we examine how to use our knowledge of the layer structure, based on the percolation theory and the variational principle, to design basis functions that will capture a global layer accurately. We find that the framework developed in this section can be used more broadly as we will see in the next chapter.

Ideally, we would like to construct special basis functions $\Psi_i, i = 1, \ldots, s$ such that $u - \sum_1^s b_i \Psi_i$ is smooth when b_i are chosen optimally. If we can do that, the smooth part can be approximated well by the bilinear functions Φ_i. Then, for $u^h = \sum_i^n a_i \Phi_i + \sum_i^s b_i \Psi_i$, we would have [79]

$$\left\| u - u^h \right\|_l \leq C\, h^{k-l} \left\| u - \sum_1^s c_i \Psi_i \right\|_k . \tag{3.32}$$

In general, it is too hard to find all these Ψ_i. It is equivalent to finding the location and behavior of all the layers in the solution exactly. For the random flow problem, we are trying to capture just the main feature of the solution which we think has the dominant contribution. We see in the numerical examples that the large errors in fact come from few sharp layers and that the use of a limited number of Ψ_i improves the result significantly.

3.6.1 Cellular Flow Case

The extra basis functions are constructed based on an insight gained from the one-dimensional example shown in Figure 3.11. We define the support of the new basis functions along the layer and then embed the correct exponential behavior. At each location, there are two bases, each with the predicted layer shape on one side, and linear function on the other. The underlying grid is uniform and the bilinear bases on this grid takes care of the smooth part of the solution.

A simple test problem is for the cellular flow case. We compute the problem (3.11) with $\delta = .25$. We see the presence of strong layers in the solution, which is shown in Figure 3.13. Where the exponential functions are present, we find that the exponential bases pick up most of the weight while linear functions contribute little. These extra functions suppress oscillations and reduce the error, as seen in Table 3.4. The method with exponential bases is less sensitive to ϵ.

In Section 3.4.3, we found that the test function with the scaling exponent $\alpha = 1/2$ had the minimum energy for the cellular flow case. To see whether that is a useful estimate for the present situation, we compute a sequence of solutions, with a set of additional basis

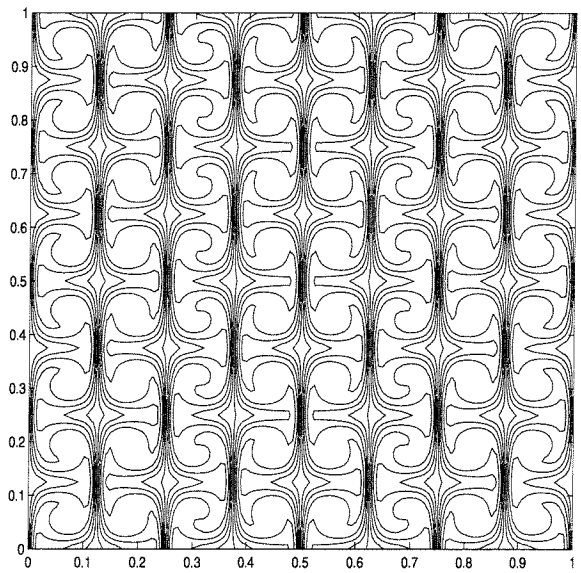

Figure 3.13: χ for the cellular flow ($\delta = .25$).

ϵ	l^2 error	l^∞ error
With extra bases		
0.001	0.001645	0.006589
0.0005	0.002460	0.010542
0.00025	0.003368	0.015862
Linear bases only		
0.001	0.001784	0.010459
0.0005	0.003957	0.020935
0.00025	0.008201	0.041641

Table 3.4: Improvement in error with exponential basis functions for the cellular flow ($n = 64, m = 16$). As before, the global coarse grid is $n \times n$. The mesh inside each element for numerical integration is $m \times m$.

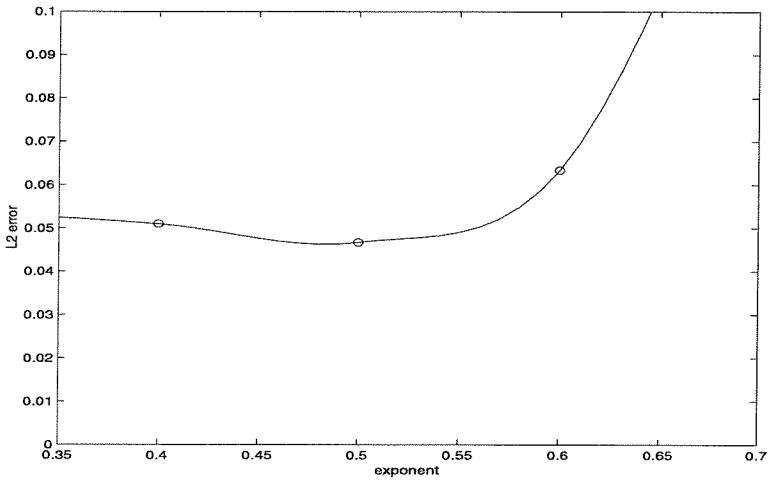

Figure 3.14: l^2 error of the cell problem as a function of the exponent α; the error is minimized with $\alpha = .5$, as was indicated by the variational integral. A cubic spline was fit using additional points (not shown).

functions whose exponential behavior is given by different α each time. In Figure 3.14, we plot the l^2 error, as a function of the exponent α. We note that indeed the minimum error is achieved when the basis functions are constructed with $\alpha \approx 1/2$. This is an important verification. It confirms that the value of α obtained through the variational principle is the value that we can use to approximate the exponential behavior. Figure 3.14 also shows that the method is robust with respect to a slight error in the estimate of α.

3.6.2 Random Flow Case

In general, the streamlines along which the layers occur are not aligned with the underlying uniform grid. The computer implementation in this case becomes more involved. For example, we need to segment evenly the separatrices on which layers occur, define the element support appropriately, and use a mapping from a quadrilateral support to a square to construct modified linear bases.

While the variational principle of the last section gives information about a dominant layer, this may not be sufficient for some situations. If δ is not small, for example, there may not be one such layer; there may be other layers we need to capture. We thus modify the method slightly. We know that the internal layers will generally occur along the separatrices

that contain the hyperbolic points. The layers occur there because that is where the flows in opposite directions meet. Therefore, we search for these hyperbolic points near the $\psi = 0$ using a searching algorithm similar to the one given in [27], and place additional extra bases functions nearby.

We study the performance of the method in the following "cell" problem. In the random flow case, the cell problem (2.6) does not appear to be useful at first because it is now over the whole domain $\mathcal{R}^2$. But since we have a convenient characterization (2.5) of σ_ϵ in terms of the cell problem solution χ, we still would like to use the cell problem. We do this by thinking of the random problem as an infinite volume limit of the periodic cell problem.

We first generate some random streamfunction, such as the one shown in Figure 3.15. Since the formulation of the cell problem requires periodic boundary conditions, we extend the original streamfunction in the first quadrant to the other quadrants by reflections $\psi(x, y) = \psi(x, -y)$ and $\psi(x, y) = \psi(-x, y)$ [33]. We then rescale the domain to $[0, 1]^2$ before solving the problem. A resolved solution for the velocity field given by the streamfunction of Figure 3.15 extended in this way is shown in Figure 3.16. To use the exponential basis functions, we lay down a mesh as shown in Figure 3.17. The boxes along the curved streamlines indicated the support of the special basis functions. Along this curved mesh, we place a set of exponential basis functions, in a manner analogous to the one-dimensional example of Section 3.5.

The results for the $\epsilon = .001$ case are shown in Table 3.5. We see that the error is decreased by 60% for both in the l^2 and l^∞ norm when exponential functions are used. Since changing the resolution in the interior of the basis function results in only small changes, most of the error is due to other weak layers not approximated by the exponential form or due to the differences in the exponential behavior on different layers.

For smaller ϵ, the difference is larger. For $\epsilon = .0005$, we see that the multigrid method that we found to be very robust does not converge in the bilinear case unless n is at least 64. For $\epsilon = .00025$, we need at least $n = 256$ for its convergence. With the exponential basis, smaller ϵ cases are handled more effectively. We see in Table 3.6 that stable solutions are obtained, although the error does increase, at the rate of roughly 40%, when ϵ is halved.

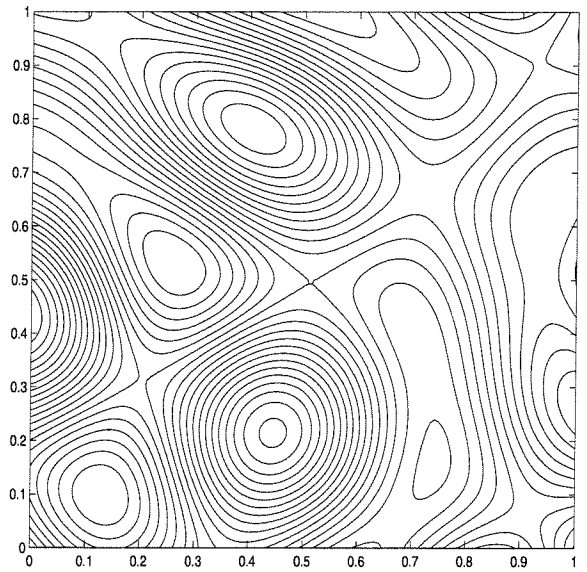

Figure 3.15: A random streamfunction for the use of exponential basis functions. This is the streamfunction for a quarter of the domain.

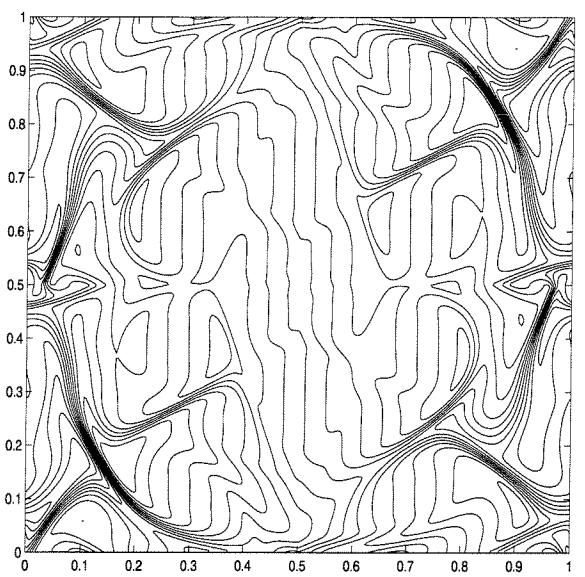

Figure 3.16: Solution to a random cell problem. The streamfunction was generated by extending the random streamfunction of Figure 3.15 to all four quadrants and then rescaling it to the unit square. The boundary conditions are now periodic.

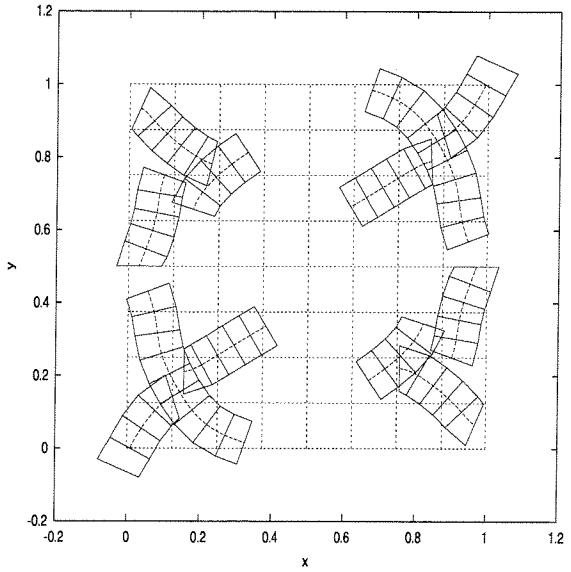

Figure 3.17: Support for the exponential basis functions along some separatrices.

	n	m	l^2 error	l^∞ error
linear	16	16	0.089987	0.28294
exponential	16	8	0.038651	0.11858
	16	16	0.037504	0.11340
	16	32	0.037553	0.11117

Table 3.5: Improvement due to the use of exponential functions ($\epsilon = .001$).

ϵ	n	m	l^2 error	l^∞ error
.001	16	16	0.037504	0.11340
.0005	16	16	0.055643	0.15914
.00025	16	32	0.080120	0.21643

Table 3.6: Dependence of the error on ϵ using the exponential basis functions.

α	l^2 error	l^∞ error
0.25	0.041468	0.14996
0.30	0.037868	0.12832
0.35	0.037504	0.11340
0.40	0.038647	0.12414
0.45	0.040918	0.13583
0.50	0.042721	0.14491

Table 3.7: Behavior of the error for varying values of α ($\epsilon = .001, n = 16, m = 16$).

It is not always possible to determine the shape of the layers exactly. The widths also vary a little on different layers. Therefore, it is desirable to have low sensitivity to small inaccuracies in the estimates for α, the exponent that determines the layer thickness given some ϵ. In Table 3.7, we show the error in the solutions for different α. We see that $\alpha \approx 0.35$ gives the least error. This value is not surprising since we expect it to be somewhere between $\alpha = .5$ of the cellular flow and $\alpha \approx .23$ of random flow in the infinite domain limit. We see in the table that as α changes little, the error increases only by a small amount. This shows that the method is tolerant to small errors in the estimate of α.

Ideally, we would like to use an efficient method such as the one developed in this section to compute the diffusivity scaling for the random flow accurately. However, in the steady-state problem, we find that the layer does not form a long smooth curve. That may be true in the time-dependent problem but still only in the long-time limit. Instead, due to the nonhomogeneous boundary condition, the solution develops more small scale features, as shown in the contour plot of Figure 3.7. We see that the layer locations are very complicated. We find that the resulting complexity of these layers is too burdensome for the method of this section. The fact that the method seems to insensitive to small errors in the layer thickness estimate works to our advantage, but the location of the layers are not clearly known. As δ gets small, there are many potential sites for these layers and it becomes difficult to account for all of them. For less complex problems, the exponential basis approach works well, as we demonstrated above with numerical examples.

There is one important difference between conventional methods and the exponential approach we consider here: because the latter is based on asymptotic theory, its performance should improve as the parameters get small. For example, if we can manage to evolve the

time-dependent problem for a long time, there should be more coherent structure in the solution, as predicted by the percolation results. Then the present method should capture more of the important features in the solution. In contrast, other methods do not take advantage of any solution structure and would deteriorate quickly for smaller parameters.

3.6.3 Solution of the Augmented Linear System

The global stiffness matrix of the conventional finite element method in two dimensions is a banded matrix with nine diagonals. With the extra basis functions, however, the stiffness matrix contains added rows and columns, as shown in (3.33).

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)\left(\begin{array}{c} x_1 \\ \hline x_2 \end{array}\right) = \left(\begin{array}{c} f \\ \hline g \end{array}\right) \tag{3.33}$$

The matrix A represents the interaction of the bilinear elements, B and C the interactions between the bilinear and the exponential, and D the interaction between the exponentials; x_1 is the weight of the bilinear elements and x_2 that of the special elements. The support of the exponential basis function does not have to be the same size as that of a bilinear element; thus one exponential basis function can interact with more than eight bilinear elements. As a result, B, C, and D in general do not have a regular pattern and must be treated as full matrices.

In solving this system, we still would like to take advantage of the banded structure of A. The following algorithm is a generalization of the Bordering algorithm in [51]. One way of combining the system

$$Ax_1 + Bx_2 = f \tag{3.34}$$

$$Cx_1 + Dx_2 = g \tag{3.35}$$

results in the form with the Schur complement,

$$\left(D - CA^{-1}B\right) x_2 = g - CA^{-1}f. \tag{3.36}$$

Based on this equation, we solve the original problem in the following manner:

i) Solve for η in $A\eta = B$.

ii) Let $L = D - C\eta$.

iii) Solve for ξ in $A\xi = f$.

iv) Let $r = g - C\xi$.

v) Solve $Lx_2 = r$.

vi) Solve $\xi - \eta x_2$.

Thus we can still use the fast multigrid solver [28] for the banded matrix A. Step (i) is the most expensive, since we need to solve the matrix problem for each column of B. However, the multigrid iteration can be supplied with a good initial guess and hence is more efficient after the first solve. Many columns of B are also zeros and do not require the iteration.

One difficulty still arises for the cell problem (2.7). Dirichlet and Neumann boundary conditions can be applied to the extended matrix (3.33) without much trouble, but when the periodic boundary condition is required for the cell problem, the banded structure of A is destroyed. We then store all the matrices in a sparse form [70] and supply the matrix-vector multiplication routine to the GMRES [13] algorithm. It is not as fast as the multigrid, but we can still obtain a good solution.

3.7 Summary

We have considered the convection-diffusion equation with a random velocity field. Because the characteristic length scale of the problem is larger than each element, methods employed

in Chapter 2 fail to capture the correct behavior. A nonlocal information is needed to construct the desired basis functions.

A solution is provided by a variational principle [31], which was obtained through a symmetrization process. We find that we can use this variational principle as a numerical method. We construct artificial test functions based on the fact that the dominant contribution to the effective diffusivity comes from an exponential layer along a long streamline near the zero level set of the streamfunction. Using these test functions in the variational integral, we are able to find the correct scaling for the cellular flow and obtain the scaling for the random flow that seems to support a theoretical prediction. This prediction is also in agreement with the direct simulations we carried out on a parallel computer.

Based on the layer thickness and location estimates thus obtained, we construct exponential basis functions and superimpose them on the uniform grid. The proper construction of the basis is not obvious and requires insights from a one-dimensional problem with a turning point. We find that with these extra basis functions, we can suppress spurious oscillations and obtain small errors with a relatively coarse grid.

Chapter 4 Elliptic Problems with High Contrast Coefficients

4.1 Introduction

In the previous chapters, we have considered the singularly perturbed elliptic equation that has the characteristics of a hyperbolic equation. Now we consider the genuinely elliptic case, but with a degeneracy coming from a large variation in the values of the coefficients. In general, the coefficient may be a matrix (a_{ij}), but here we will deal with the isotropic case, $a_{ij} = a\,\delta_{ij}$. The equation is

$$-\nabla \cdot (a(x,y)\nabla u) = f, \quad \Omega \in \mathcal{R}^2, \tag{4.1}$$

with $a(x,y) \geq a_0 > 0$. The finite element formulation, as before, is to find $u^h \in U^h$ such that

$$\int_\Omega a\nabla u^h \cdot \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x}, \quad \forall v \in V^h, \tag{4.2}$$

where U^h and V^h are some finite element spaces containing trial and test functions, respectively.

Elliptic equations [35] can be found in many areas of science and engineering, such as flow through porous media and conductivity in composite materials, and they often have large variations in their coefficients. For instance, $a(x,y)$ may represent permeability or electrical conductivity, which can often vary wildly in different regions. We refer to these as the "high contrast problems," where the contrast is defined as $\max a(x,y)/\min a(x,y)$.

While conventional numerical methods work well for smooth and slowly-varying coefficients, they need to be modified in other cases. For highly oscillatory coefficients, for example, the

multiscale FEM was developed and tested extensively in [44, 43]. In this chapter, we deal with the modifications for high contrast problems, as well as problems with singularities arising from discontinuities in the coefficients [79, 9, 1]. We are particularly interested in the case in which the coefficient is discontinuous and the contrast is large.

We again employ a finite element method with special basis functions in this chapter. In the convection-diffusion problems of the previous chapters, the dominance of the convective effects over the diffusive effects, as expressed by a large Peclet number, created sharp features in the solution. Special exponential basis functions were introduced in place of mesh refinements. In the first problem we consider in this chapter, the large variance in the coefficient also causes sharp features, for which special basis functions will be introduced. In general, it is difficult to find the terms of the asymptotic expansion for the solution explicitly, but this can be done in some high contrast problems [17]. We utilize this result in devising additional basis functions.

In Section 4.3, we consider the case of discontinuous coefficients. This creates difficulties of varying degrees depending on the smoothness of interface and the contrast. It is well known that the convergence rate of the standard numerical methods is degraded in most cases, and it can even be arbitrarily slow in the case of severe singularities. Even the adaptive mesh refinement schemes may become too burdensome very quickly for some of the singularities. We start with a review of the properties of the singularities and some available methods to deal with them. Most methods require that the structure of the singularity is known beforehand, often by solving a related eigenvalue problem. Only after the leading behavior is known can those methods perform efficiently.

We turn to the infinite element method to remedy the situation. Using the special self-similar grid placed around the singularity, the method is designed to offer essentially infinite resolution at the singularity. The number of unknowns meanwhile stays proportional to the number of mesh points on the boundary of the element. If an element has m points on each side, the matrix problems involved are for $4m$ unknowns, not for m^2 unknowns as is the case with finite elements. We use this technique to construct the basis if that element covers the singularity. We can use bilinear elements away from the singularity. We apply this infinite element approach to interface problems with corner singularities and the

"checkerboard" singularities. Through numerical experiments, we find that the method we propose is efficient and robust, particularly with regard to the high contrast.

4.2 High Contrast Problems

4.2.1 Description of the Problem

In the porous media problems, the coefficient may represent conductivity or permeability; one can easily see that this quantity can be very large or small depending on the medium. It is not unusual for the contrast to be on the order of 10^4. We deal with this type of coefficients in this section, but with the assumption that the coefficients are still continuous.

The important result in this case is the flow channeling: the flow tends to concentrate in narrow channels in between the highly conductive or resistive inclusions of the media. This was studied by Batchelor and O'Brian in [14]; other references are in [17]. This means that these channels are the key features that need to be computed accurately. Away from these channels, the resolution needs not be high.

Previous studies, e.g. [52, 18, 55], have revealed that this concentration of the flow can be modeled well by the conductivity of the form

$$a(x,y) = a_0 \exp\left(-S(x,y)/\epsilon^2\right), \tag{4.3}$$

where $S(x,y)$ is a smooth function with a saddle point and $\epsilon \ll 1$ is the parameter that determines the contrast. This has been shown to be a good model even when the true conductivity is not as smooth as represented by $S(x,y)$ [55].

In [17], Borcea and Papanicolaou used the asymptotic properties of this model to extract the behavior of the solution at the channels. This information was then incorporated in the finite difference method to make a hybrid method. Given the extra information in the form of an asymptotic behavior, one should be able to design a more efficient method. Their numerical results confirm that the error of the hybrid method is smaller compared to the

standard finite difference scheme and is often comparable to that of the more expensive mesh refinement schemes.

Our approach is similar to that of Chapters 2 and 3, where we exploited the exponential behavior of the solution layers to supplement the finite element space with exponential basis functions. This is a natural setting in which to incorporate the asymptotic results, even though the layer may not be exponential. In fact, we can avoid extra procedures needed in [17], such as computing the fluxes and matching them across boundaries to link to the finite difference method. In the finite element setting, the best combination of basis functions are chosen automatically and the channels can be placed in arbitrary locations without much additional work.

4.2.2 Asymptotic Behavior

We first discuss the asymptotic result, as reviewed in [17]. In the small neighborhood of the saddle point $\mathbf{x}_s$, the conductivity behaves as

$$a(x,y) = a_0 \exp\left(-S(x,y)/\epsilon^2\right) \approx a_0 \exp\left(\frac{k_1(x-x_s)^2}{2\epsilon^2} - \frac{k_2(y-y_s)^2}{2\epsilon^2}\right) \tag{4.4}$$

with some positive constants k_1 and k_2. The saddle could be pointing in any direction, but we have chosen it to be in the x direction here for simplicity. Then, in that neighborhood, the solution can be obtained by the matched asymptotic expansion [54, 41] to be

$$\phi(\mathbf{x}) \approx -C \operatorname{erf}\left[\sqrt{\frac{k_1}{2\epsilon^2}}(x-x_s)\right], \tag{4.5}$$

where $\operatorname{erf}(x)$ is the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt.$$

From this, we can get the flux using Darcy's law:

$$\mathbf{j} = -a\nabla\phi \approx a(\mathbf{x}_s)\; C \sqrt{\frac{2k_1}{\pi\epsilon}} \exp\left(-\frac{k_2(y-y_s)^2}{2\epsilon^2}\right) \mathbf{e}_x. \tag{4.6}$$

Since this flux is a narrow Gaussian centered at $\mathbf{x}_s$, we see that the flow concentration is high near the saddle point.

In general, there may be many channels in the domain. In that case, we can model each of them separately as above. In [18], the case of many channels was modeled as a network of resistors.

To have the desired saddle point behavior (4.4), $S(x, y)$ must be chosen carefully. One choice [17] is

$$S(x, y) = S_0 + M(\xi)M(\eta)\ \sin(\gamma\xi)\sin(\gamma\eta),$$

where S_0 is the background conductivity, M is the mollifier

$$M(x) = \begin{cases} 0, & x < -\pi/\gamma \\ \sin^3\left(\frac{\pi}{2}\left(x + \frac{\pi}{\gamma}\right)/d\right), & -\pi/\gamma < x < -\pi/\gamma + d \\ 1, & -\pi/\gamma + d < x < \pi/\gamma - d \\ -\sin^3\left(\frac{\pi}{2}\left(x - \frac{\pi}{\gamma}\right)/d\right), & \pi/\gamma - d < x < \pi/\gamma \\ 0, & \pi/\gamma < x \end{cases} \tag{4.7}$$

and ξ, η are the rotated coordinates

$$\begin{aligned} \xi &= \left(x - \frac{1}{2}\right)\cos\theta + \left(y - \frac{1}{2}\right)\sin\theta, \\ \eta &= \left(x - \frac{1}{2}\right)\sin\theta + \left(y - \frac{1}{2}\right)\cos\theta. \end{aligned}$$

We take $\gamma = 2\sqrt{2}\pi$ for the computational domain of the unit square and $\theta = \pi/4$ to have the saddle orientation match that of (4.4). We plot this $S(x, y)$ in Figure 4.1.

4.2.3 A Special Basis Function

In Figure 4.2, we plot the asymptotic result (4.5) for various values of $\sqrt{k_1/(2\epsilon^2)}$. The error function is an integral of an exponential function, but it still has a sharp layer for small ϵ. The asymptotics gives an excellent approximation, even for moderate values of ϵ, to the

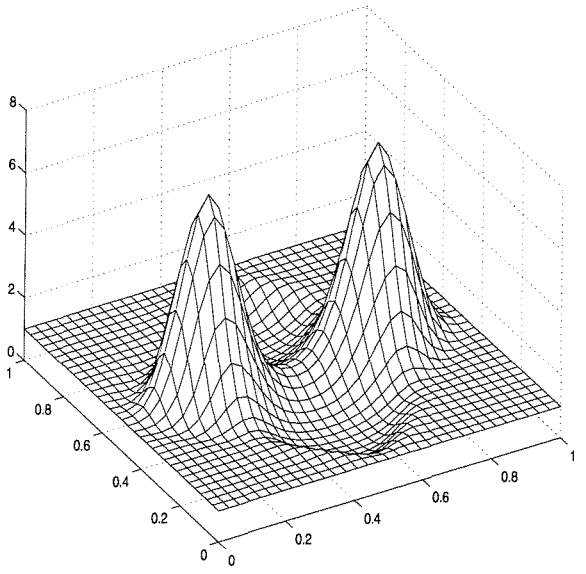

Figure 4.1: A high contrast coefficient ($\epsilon^2 = .5, d = \sqrt{2}/8$).

shape of the solution.

We now construct the special basis function using the asymptotic expression. The procedure is similar to that of Chapter 3. We define a set of basis functions, in addition to the bilinear ones, that have the error function behavior of (4.5) on one side while linear on the other. We lay down a series of these bases along the channel. Depending on the location of the saddle, the support of these bases may or may not coincide with the underlying grid. But this has been worked out in Chapter 3 and presents no serious difficulty. An example of this basis is shown in Figure 4.3.

4.2.4 Numerical Examples

We consider a sample problem with Neumann boundary conditions and driven by the forcing function of the form

$$f(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_1) - \delta(\mathbf{x} - \mathbf{x}_2), \tag{4.8}$$

where $\mathbf{x}_1$ is the point source and $\mathbf{x}_2$ is the sink. As usual, we approximate the delta functions with regularized Gaussians. A resolved solution is plotted in Figure 4.4. We see the flow

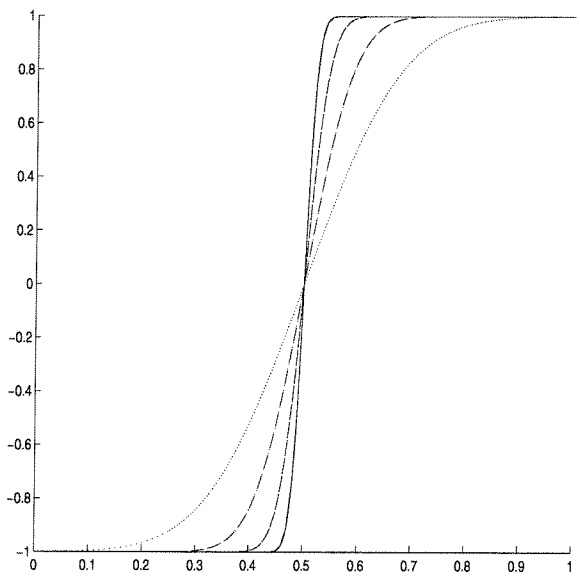

Figure 4.2: The asymptotic behavior (4.5) of the channel, with $x_s = .5$. ϕ is plotted on the y-axis ($\sqrt{k_1/(2\epsilon^2)} = .2, .1, .05, .025$ and $C = -1$).

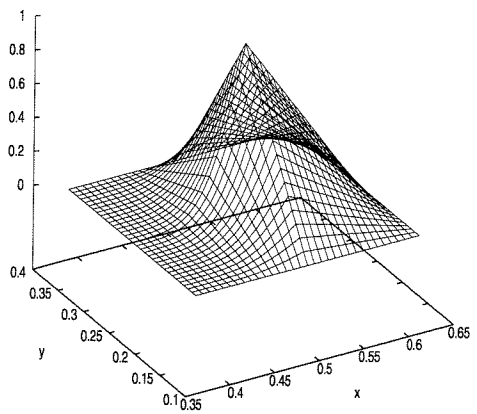

Figure 4.3: A basis function using the asymptotic result ($c_1 \text{erf}((x - x_s)/0.05) + c_2$ on one side and linear on the other).

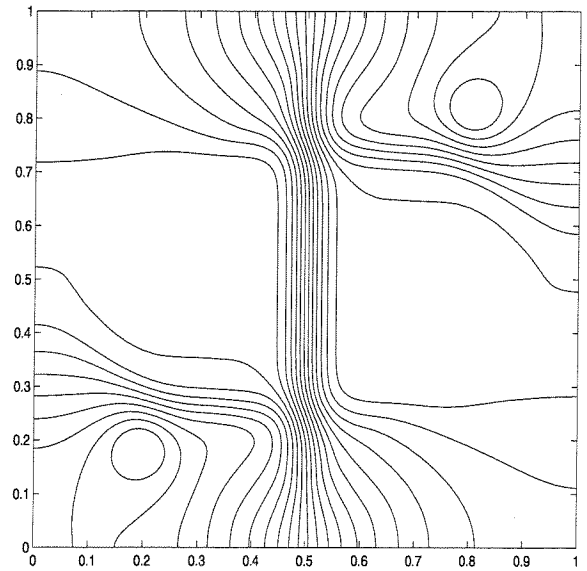

Figure 4.4: A contour plot of the solution with high contrast conductivity; there is a narrow channel through the saddle point ($\epsilon^2 = .15, \mathbf{x}_1 = (.2, .2), \mathbf{x}_2 = (.8, .8)$).

n	ϵ^2	contrast	l^∞ relative error	l^∞ rate
16	0.2166168	1E4	1.12%	
16	0.1732940	1E5	2.80%	2.50
16	0.1444116	1E6	3.83%	1.37
16	0.1237813	1E7	4.56%	1.19
16	0.1083087	1E8	5.49%	1.20
16	0.0962744	1E9	6.43%	1.17

Table 4.1: Relative error in l^∞ for different contrasts.

concentrates along the channel, as dictated by the saddle point behavior. Other features are smooth and can be approximated by bilinear functions.

With the asymptotic basis functions along the channel, most of the weight is picked up by the extra bases and there is a large improvement in the solution. For $\epsilon = .15, n = 16$, the relative error in l^2 or l^∞ norms with only the linear functions is 17.1% while it is 3.6% with the asymptotic bases. This is about 80% reduction in the error. In addition to this large reduction, the error is not sensitive to the change in contrast. We show this in Table 4.1. We see that the relative error increases only slightly for a tenfold increase in contrast, unlike conventional methods.

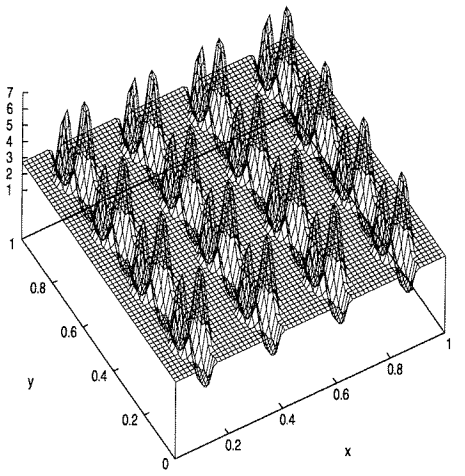

Figure 4.5: Coefficient for the multiple channel case ($\epsilon = 1$).

n	ϵ^2	contrast	relative error in l^∞
16	0.1732940	1E5	2.93%
16	0.1444116	1E6	3.70%
16	0.1237813	1E7	4.52%
16	0.1083087	1E8	8.28%

Table 4.2: Relative error in l^∞ for multiple channels.

Our method performs similarly for different configurations of channels. We simply need to make sure that the asymptotic basis functions are placed along the predicted channel paths. The cases in which the channels occurs in complicated paths are handled easily by the implementation developed in the previous chapter. We can generate a more complicated structure, for example, by extending a single channel case periodically. A coefficient with 16 saddle points is shown in Figure 4.5. The numerical results are presented in Table 4.2 for various values of ϵ. Again, the error is not too sensitive to the contrast. The error is larger if we use bilinear elements only. For example, for 10^5 contrast, using bilinear elements only give relative errors of 16%, as opposed to 2.9% for combining them with asymptotic bases.

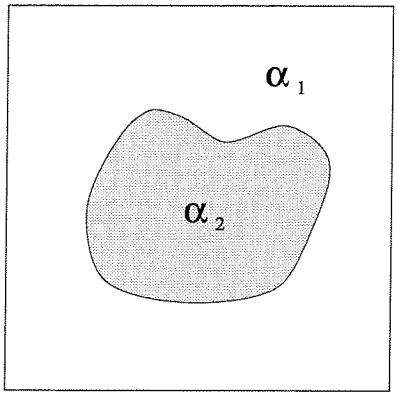

Figure 4.6: The simplest problem with smooth interface.

4.3 Infinite Elements for Discontinuous Coefficients

When the coefficients are discontinuous, we require the governing equation (4.1) to hold in each domain and also that the solution and the normal flux be continuous across the interface.

Suppose we have

$$a(x,y) = \begin{cases} \alpha_1 & \text{in } \Omega_1, \\ \alpha_2 & \text{in } \Omega_2, \end{cases}$$

where Ω is a convex polygon in $\mathcal{R}^2$, $\Omega_1 \subset \Omega$ is an open domain with boundary $\Gamma = \partial\Omega_1 \subset \Omega$, and $\Omega_2 = \Omega \setminus \Omega_1$. An example is shown in Figure 4.6 in the case of a smooth interface.

The continuity and the jump conditions are

$$u_1|_{\partial\Omega_1} = u_2|_{\partial\Omega_1}, \qquad \alpha_1 \frac{\partial u_1}{\partial n}\bigg|_{\partial\Omega_1} = \alpha_2 \frac{\partial u_2}{\partial n}\bigg|_{\partial\Omega_1}, \tag{4.9}$$

where n is the outward normal to the interface $\partial\Omega_1$. We now discuss the singularities that arise in this context, depending on the smoothness of the interface $\partial\Omega_1$. In the next section, we introduce the infinite element method, followed by a section on how it can be constructed as a basis function. We present numerical results in Section 4.3.4.

4.3.1 Behavior at the Singularities

Smooth Interfaces

When the interface is smooth, as in Figure 4.6, the singularity is not severe. A smooth interface implies that the solutions are smooth inside each region, although the global regularity may be lower. Many methods have been developed for this case and they work well, at least for moderately large contrast. A method that is the simplest conceptually but nontrivial in implementation is aligning the grid with the discontinuity. In [23], it is proved that if the boundary is at least C^2, this finite element method converges nearly in the same optimal way, in both the L^2 and energy norms, as in the problems without interfaces.

In [58], Leveque and Li develop the Immersed Interface Method. They use a uniform rectangular grid, but they develop different finite difference formulas near the discontinuities, first expanding around the discontinuity in Taylor series and then using the jump condition to substitute terms in the expansion. One problem with this method is that while the continuous problem is self-adjoint and uniformly elliptic, the linear system may be non-symmetric and not positive definite. When the contrast is large, the system is also highly ill-conditioned and hence difficult to solve accurately. This method also cannot handle interfaces that are not smooth.

In the finite element formulation, Babuška and Osborn in [10] give one way of modifying the trial functions without aligning the mesh with the interface, which is assumed to be a convex polygon. Let $\partial\Omega$ be the interface, $\mathcal{T}_h$ a quasi-uniform triangulation, and n and t are the normal and tangent to $\partial\Omega$. For an element $T \in \mathcal{T}_h$, the basis function ϕ is linear if $T \cap \partial\Omega = \emptyset$; if $T \cap \partial\Omega \neq \emptyset$, it satisfies the following conditions:

i) $\phi_1 = \phi|_{T\cap\Omega_1}$ and $\phi_2 = \phi|_{T\cap\Omega_2}$ are linear,

ii) $u_1|_\Gamma = u_2|_\Gamma$, where $\Gamma = \partial\Omega \cap T$,

iii) $(\partial u_1/\partial t)|_\Gamma = (\partial u_2/\partial t)|_\Gamma$,

iv) $\alpha_1(\partial u_1/\partial n)|_\Gamma = \alpha_2(\partial u_2/\partial n)|_\Gamma$.

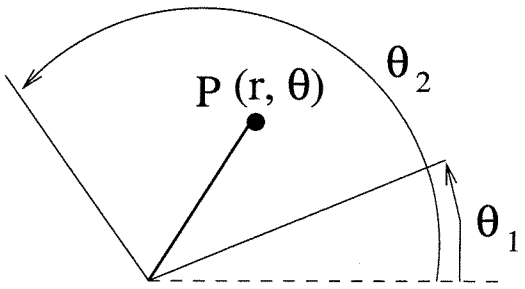

Figure 4.7: A corner with interfaces at θ_1 and θ_2.

The following energy norm estimate is proved [10]:

$$\|u - u_h\|_{H_{1,h}} \leq Ch\|f\|_{L^2},$$

where

$$\|v\|^2_{H_{1,h}} = \int_\Omega v^2 \, dx \, dy + \sum_T \left(\int_{T \cap \Omega_1} |\nabla v|^2 \, dx \, dy + \int_{T \cap \Omega_2} |\nabla v|^2 \, dx \, dy \right).$$

However, when the contrast gets large, the stiffness matrix again becomes ill-conditioned and it is difficult to obtain an accurate solution.

Corner Singularities

The singularity caused by a corner in the interface is more difficult to handle. The severity of the singularity depends on the shape of the corner and the contrast in the coefficient, as revealed by the eigenvalues of the Sturm-Liouville problem described below. When f is sufficiently smooth on the boundary, the solution in the neighborhood of the singularity can be expanded [38] as

$$u(r, \theta) = \sum_{i=0}^{\infty} \gamma_i \, r^{\nu_i} \, \Theta_i(\theta),$$

where r and θ are as shown in Figure 4.7.

The exponents ν_i are the eigenvalues of the problem

$$\left(a\Theta'\right)' + \nu^2 \Theta = 0,$$

with the associated eigenfunctions Θ_i. These ν_i determine how badly the singularity behaves. We see that if $\nu_0 = .1$, for example, $r^{.1}$ is very steep at the origin and requires many points to resolve.

If we denote the angles that the interfaces make at the singularity by θ_i, then the jump conditions (4.9) become

$$\Theta(\theta_i^+) = \Theta(\theta_i^-), \tag{4.10}$$

$$a(\theta_i^+)\Theta'(\theta_i^+) = a(\theta_i^-)\Theta'(\theta_i^-). \tag{4.11}$$

The eigenfunctions are orthogonal,

$$\int_0^{2\pi} a(\theta)\,\Theta_i(\theta)\,\Theta_j(\theta)\,d\theta = \delta_{ij},$$

and we can solve for the constants γ_i [79]. For simple problems, this process can be carried out and we can obtain exact formulas for eigenvalues and eigenfunctions. But in general, this must be solved numerically. Solving this problem is the first step in many numerical methods [79]. Without the knowledge of the singularities, these methods cannot be efficient.

Another class of methods for handling the corner singularity is similar to the "auxiliary mapping" by Babuška and Oh [66]. The idea is to use a local transformation to map a region around the singularity to a new domain. For a simple geometry, a conformal mapping of type $z = \xi^\alpha$ can be used. However, when the domain or the interface is more complicated, some other appropriate mapping must be found somehow. Moreover, the nature of the singularity must be obtained separately in order to determine α.

In the new method we introduce in the next section, there is no need to solve this eigenvalue problem. The local grid will have sufficient resolution, even if ν_i are small.

Checkerboard Singularities

The most difficult case is that of the "checkerboard" configuration, shown in Figure 4.8 with the singularity at the center. Sometimes it is referred to as the "four-corner juncture"

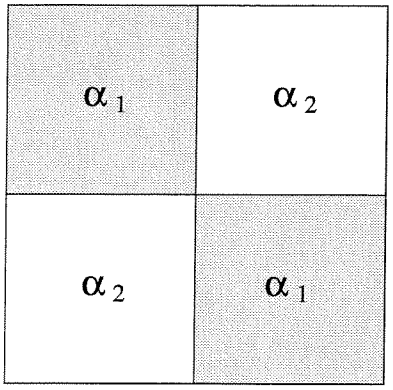

Figure 4.8: The checkerboard singularity at the center.

problem.

In general, the four regions may have four different conductivities, $\alpha_i, i = 1, ..., 4$. In the neighborhood of the singularity, the analytic solution behaves as [75]

$$u(r, \theta) = w_1(\theta) r^{\gamma} + w_2(\theta) r^{2-\gamma} + \mathcal{O}(r^2), \tag{4.12}$$

where $\gamma = \gamma(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $0 \leq \gamma \leq 1$. When three of the four regions have the same α, as in a corner of a rectangle, we get $\gamma = 2/3$. For our numerical example, we consider the case when $\alpha_1 = \alpha_3 = 1$ and $\alpha_2 = \alpha_4 = \alpha$. In that case, $\gamma = \sqrt{\alpha}$ [75]. So when the contrast is 10000, the leading order behavior at the singularity is $r^{0.01}$. We note that as $\alpha \to 0$, we also have $\gamma \to 0$. In fact, in the limit of $\alpha \to 0$, the solution becomes discontinuous at the singularity. This problem has been studied before, often as an example of when a method fails.

Since the singularity is local, we can use regular finite elements in the smooth region away from the singularities. Near the singularity, we employ the "infinite elements." The infinite element method is an elegant technique that has the advantages of having essentially infinite mesh refinement at the singularity. We introduce this method in the next section.

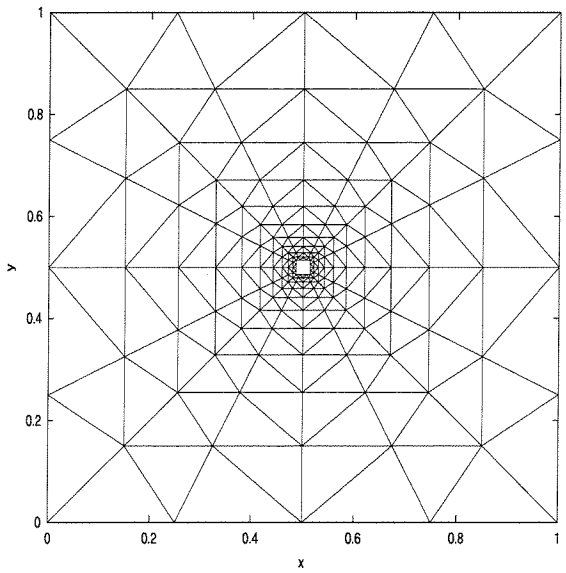

Figure 4.9: An example of the infinite element grid; further refinements at the center are not shown in the picture.

4.3.2 Infinite Element Method

First appeared in early 1970s, the infinite element method has been most useful in infinite domain problems and singularity problems, for reasons that will become obvious. Its many properties and some applications are described in [82] and the references therein. The main idea is to lay down a grid with certain similarity structure and then utilize the self-similarity in reducing the size of the problem.

For example, consider a square domain with a singularity at the center. We lay down a grid that looks like the one in Figure 4.9. The dimensions of the elements decrease geometrically by a factor of $0 < \xi < 1$ as they get closer to the center. Close to the center, the elements get arbitrarily small and there is essentially an infinite number of refinements. We call the region between two successive nested squares a "layer" (see Figure 4.10). In general, the element can be any polygon and the singularity does not have to be located at the center. The only requirement is that the similarity is preserved between the layers and in the triangulation of each layer. In partitioning inside a layer, we do it in such a way as to avoid small angles at the element corners.

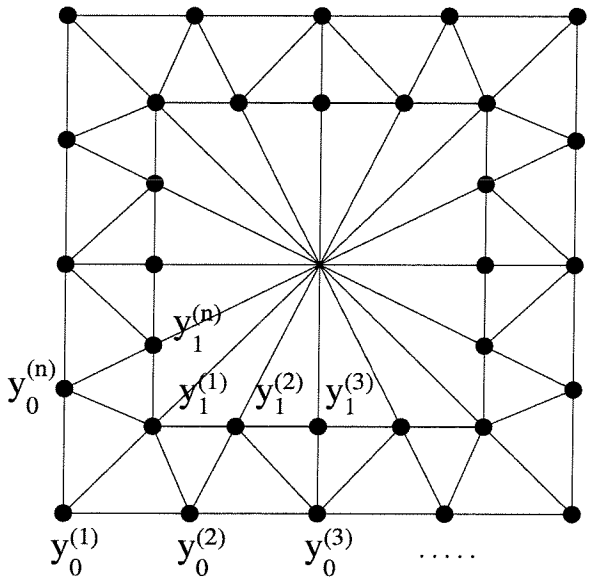

Figure 4.10: The node numbering scheme for the first layer.

For the following discussion, we label the nodes by letting $y_k^{(i)}, (k = 0, ..., \infty; i = 1, ..., n)$ be the i^{th} point around the center on the k^{th} layer. This is illustrated in Figure 4.10. Now, to create the global stiffness matrix, we add up the local stiffness matrix for each element to get a "layer stiffness matrix." It can be written as

$$\begin{pmatrix} K & -A^T \\ -A & K' \end{pmatrix}, \tag{4.13}$$

with K and K' being positive definite. This is a $2n \times 2n$ matrix that relates $y_k^{(i)}, (i = 1, ..., n)$ and $y_{k+1}^{(i)}, (i = 1, ..., n)$ for any k. The key observation is that this matrix is the same for all layers due to the self-similarity of the grid.

For the global stiffness matrix, we put the layer matrices together and obtain

$$\begin{pmatrix} K & -A^T & & \\ -A & K'+K & -A^T & \\ & -A & K'+K & \ddots \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} b_0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}. \tag{4.14}$$

This is an infinite set of equations. b_0 comes from the boundary condition. In order to solve

this system, we first assume that there is a matrix X, called the *transfer* matrix, such that

$$y_{k+1} = Xy_k. \tag{4.15}$$

The existence of this matrix is proved in [82]. This transfer matrix plays a crucial role: given the boundary condition y_0, we can compute the values at all the interior grid points $y_k, (k = 1, ..., \infty)$ from X through (4.15). This means all the information about the singularity is contained in X.

Notice that we can write each block line after the first in (4.14) as

$$-Ay_k + (K' + K)y_{k+1} - A^T y_{k+2} = 0.$$

With $y_{k+1} = Xy_k$ and $y_{k+2} = X^2 y_k$, this becomes

$$(-A + (K' + K)X - A^T X^2)y_k = 0.$$

We can then find X by solving

$$-A + (K' + K)X - A^T X^2 = 0. \tag{4.16}$$

This is not a trivial process. One way to solve it is to rewrite this problem to an eigenvalue problem of the form

$$\begin{pmatrix} K' + K' & -A \\ I & 0 \end{pmatrix} \begin{pmatrix} y_k \\ y_{k-1} \end{pmatrix} = \begin{pmatrix} A^T & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} y_{k+1} \\ y_k \end{pmatrix}. \tag{4.17}$$

We let λ be the eigenvalues and g the corresponding eigenvectors of X, i.e.,

$$Xg = \lambda g,$$

and (4.17) can be written as

$$\begin{pmatrix} K' + K' & -A \\ I & 0 \end{pmatrix} \begin{pmatrix} \lambda g \\ g \end{pmatrix} = \lambda \begin{pmatrix} A^T & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \lambda g \\ g \end{pmatrix}. \tag{4.18}$$

We can now find the eigenvalues λ from $\det(R_1 - \lambda R_2) = 0$, where R_1 and R_2 are the matrices on the left and right sides respectively in (4.18). Once λ and g are found, we can find X by $X = T\Lambda T^{-1}$, where T has the eigenvectors g as columns and Λ contains the eigenvalues λ along the diagonal.

While this procedure makes sense, it is rather cumbersome. Fortunately, the following iterative method can be used instead [38].

We generate a sequence of matrices K_i, K_i', and A_i by

$$K_i = K_{i-1} - A_{i-1}^T (K_{i-1} + K_{i-1}')^{-1} A_{i-1}, \tag{4.19}$$

$$K_i' = K_{i-1}' - A_{i-1} (K_{i-1} + K_{i-1}')^{-1} A_{i-1}, \tag{4.20}$$

$$A_i = A_{i-1} (K_{i-1} + K_{i-1}')^{-1} A_{i-1}, \tag{4.21}$$

$$K_z^i = K_i - \frac{A_i^T g g^T A_i}{g^T K_i' g}, \tag{4.22}$$

with $K_0 = K, K_0' = K'$ and $A_0 = A$ from (4.13). In [82], it is proven that the sequence of K_z^i converges:

$$\lim_{i \to \infty} K_z^i = K_z.$$

With K_z, we can then compute X using [82]

$$X = (K_z + K')^{-1} A. \tag{4.23}$$

The motivation for this procedure is the minimization of energy and is described in [82]. The most expensive part of the algorithm is inverting the $2n \times 2n$ matrix $(K_{i-1} + K_{i-1}')$ at each iteration; but for relatively small n, e.g., $n = 16, 32, 64$, the algorithm is still fast. Using this algorithm, we define $X^{(k)} \equiv (K_z^i + K_i')^{-1} A_i$ and iterate until $X^{(k)}$ converges to some specified tolerance. For the problems we discuss in the next section, the number of

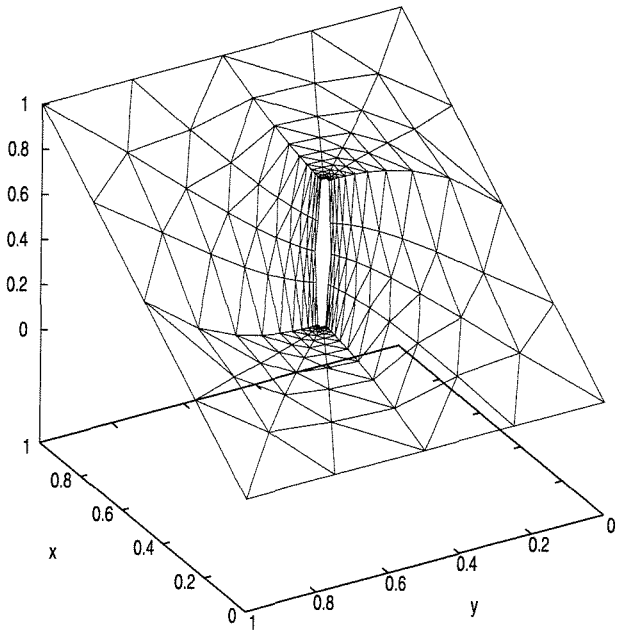

Figure 4.11: The infinite element solution for the checkerboard singularity at the center.

iterations till $\|X^{(k)} - X^{(k+1)}\| < 10^{-10}$ is fewer than 12 and is not sensitive to the severity of the singularity.

As a sample problem, we try the checkerboard configuration of Figure 4.8. We set the boundary condition to be $u(x, y) = x$. The infinite element solution is plotted in Figure 4.11, where we have computed the interior values using the transfer matrix. Only the first 10 layers were plotted in the figure, but we can clearly see that there is a singularity at the center that would be difficult to resolve by any other numerical method.

4.3.3 Infinite Element as a Basis Function

In order to use the infinite element technique for the construction of basis functions, the procedure is basically the same. The difference is that we need to impose appropriate boundary conditions for each basis. We employ the infinite elements only at the singularities; away from the singularities, we can use the standard bilinear elements. In the uniform

grid, the only requirement is that the singular part of the solution is contained in one element, which is not difficult to do since the worst singular behavior is confined to a local neighborhood.

Once we compute all the basis functions, we need to put them together into the global stiffness matrix. The integration over the linear elements can be done with the two-dimensional trapezoidal rule as usual. But over an infinite element, it is not immediately clear how to sum up the contribution of all the layers. The shape of the local elements are different and there are infinite many of them. Fortunately, we find that a simple expression can be obtained just in terms of the boundary condition and the transfer matrix, as described in the following.

We start by writing the integration over an element K as a sum over the layers,

$$\int_K \alpha \nabla \phi_j \cdot \nabla \phi_i \, d\mathbf{x} = \sum_{\substack{\text{layers} \\ k=1,\ldots,\infty}} \int_{k^{th} \text{ layer}} a \nabla \phi_i^{(k)} \cdot \nabla \phi_j^{(k)} \, d\mathbf{x}. \tag{4.24}$$

But within each layer, we can also write

$$\phi_i^{(k)} = \sum_{l=1}^{2n} \alpha_l^{(k)} \, \eta_l^{(k)}, \qquad \phi_j^{(k)} = \sum_{m=1}^{2n} \beta_m^{(k)} \, \zeta_m^{(k)},$$

where $\phi_i^{(k)}, \phi_j^{(k)}$ are the infinite element solution in the k^{th} layer, and $\eta_l^{(k)}, \zeta_m^{(k)}$ are the linear elements within the layer.

Substituting this in (4.24), we have

$$\begin{aligned} \int_{k^{th} \text{ layer}} a \nabla \phi_i^{(k)} \cdot \nabla \phi_j^{(k)} \, d\mathbf{x} &= \int_{k^{th} \text{ layer}} a \nabla \left(\sum_l \alpha_l^{(k)} \, \eta_l^{(k)} \right) \cdot \nabla \left(\sum_m \beta_m^{(k)} \, \zeta_m^{(k)} \right) d\mathbf{x} \\ &= \sum_l \sum_m \alpha_l^{(k)} \beta_m^{(k)} \left(\int_{k^{th} \text{ layer}} a \nabla \eta_l^{(k)} \cdot \nabla \zeta_m^{(k)} \, d\mathbf{x} \right). \end{aligned}$$

But we notice that each integral here is just an element of the local stiffness matrix for the infinite element basis. This has been computed already and is stored in K, K', and A in

(4.13). Here, $\alpha_l^{(k)}, (l = 1, ..., 2n)$ contains the grid point values along the outer perimeter $(l = 1, ..., n)$, and the inner perimeter $(l = n + 1, ..., 2n)$ for each layer K.

We now simplify this notation for the subsequent discussion by noticing that $\alpha_l^{(k)}$ for $l = n + 1, ..., 2n$ are the same as $\alpha_l^{(k+1)}$ with $l = 1, ..., n$; in other words, the outer perimeter points of one layer are the inner perimeter points of the next layer. We let $\tilde{\alpha}_l^{(k)}$, a vector of size n, denote only the outer points. This allows us to write

$$\alpha^{(k)} = \begin{pmatrix} \tilde{\alpha}^{(k)} \\ \tilde{\alpha}^{(k+1)} \end{pmatrix}, \qquad \beta^{(k)} = \begin{pmatrix} \tilde{\beta}^{(k)} \\ \tilde{\beta}^{(k+1)} \end{pmatrix}.$$

Now we sum over each layer. For the first layer,

$$\begin{pmatrix} \tilde{\alpha}^{(1)} & \tilde{\alpha}^{(2)} \end{pmatrix} \begin{pmatrix} K & -A^T \\ -A & K' \end{pmatrix} \begin{pmatrix} \tilde{\beta}^{(1)} \\ \tilde{\beta}^{(2)} \end{pmatrix}$$
$$= \tilde{\alpha}^{(1)T} \left(K\tilde{\beta}^{(1)} - A^T\tilde{\beta}^{(2)} \right) + \tilde{\alpha}^{(2)T} \left(-A\tilde{\beta}^{(1)} + K'\tilde{\beta}^{(2)} \right). \quad (4.25)$$

For the second layer,

$$\begin{pmatrix} \tilde{\alpha}^{(2)} & \tilde{\alpha}^{(3)} \end{pmatrix} \begin{pmatrix} K & -A^T \\ -A & K' \end{pmatrix} \begin{pmatrix} \tilde{\beta}^{(2)} \\ \tilde{\beta}^{(3)} \end{pmatrix}$$
$$= \tilde{\alpha}^{(2)T} \left(K\tilde{\beta}^{(2)} - A^T\tilde{\beta}^{(3)} \right) + \tilde{\alpha}^{(3)T} \left(-A\tilde{\beta}^{(2)} + K'\tilde{\beta}^{(3)} \right). \quad (4.26)$$

From this, we can easily see that

$$\begin{aligned}
&\sum_{\substack{\text{layers}\\ k=1,\ldots,\infty}} \int_{(k)} a\nabla\phi_i^{(k)} \cdot \nabla\phi_j^{(k)}\, d\mathbf{x} \\
&= \sum_{k=1,\ldots,\infty} \left[\tilde{\alpha}^{(k)T}\left(K\tilde{\beta}^{(k)} - A^T\tilde{\beta}^{(k+1)}\right) + \tilde{\alpha}^{(k+1)T}\left(-A\tilde{\beta}^{(k)} + K'\tilde{\beta}^{(k+1)}\right)\right] \\
&= \tilde{\alpha}^{(1)T}\left(K\tilde{\beta}^{(1)} - A^T\tilde{\beta}^{(2)}\right) \\
&\quad + \sum_{k=2,\ldots,\infty} \left[\tilde{\alpha}^{(k)T}\left(-A\tilde{\beta}^{(k-1)} + K'\tilde{\beta}^{(k)}\right) + \tilde{\alpha}^{(k)T}\left(K\tilde{\beta}^{(k)} - A^T\tilde{\beta}^{(k+1)}\right)\right] \\
&\text{with} \quad \beta^{(k)} = X\beta^{(k-1)} \quad \text{and} \quad \beta^{(k+1)} = X^2\beta^{(k-1)}, \\
&= \tilde{\alpha}^{(1)T}\left(K - A^T X\right)\tilde{\beta}^{(1)} \\
&\quad + \sum_{k=2,\ldots,\infty} \left\{\tilde{\alpha}^{(k)T}\left[-A + \left(K + K'\right)X + X^2\right]\tilde{\beta}^{(k-1)}\right\}.
\end{aligned}$$

But we had determined that $-A + (K + K')\,X + X^2 = 0$ previously in (4.16). Only the first term then survives, and we have the simple expression:

$$\int_K a\nabla\phi_i^{(k)} \cdot \nabla\phi_j^{(k)}\, d\mathbf{x} = \tilde{\alpha}^{(1)T}\left(K - A^T X\right)\tilde{\beta}^{(1)}, \tag{4.27}$$

where $\tilde{\alpha}^{(1)}$ is the boundary condition for ϕ_i and $\tilde{\beta}^{(1)}$ for ϕ_j. Therefore, the sum over an element has been reduced to multiplying the boundary condition by some matrix. This is not unexpected since the transfer matrix X should contain all the interior information.

As in the multiscale FEM case, the correct boundary condition for the infinite element basis function is not known. But in this case, good results are obtained when the reduced one-dimensional problem is used on the boundary. This is simple for a piecewise constant coefficient because an exact solution can be obtained easily. For

$$(a(x)u')' = 0, \quad u(0) = 0, u(1) = 1,$$

with

$$a(x) = \begin{cases} a_1 & \text{if } 0 < x < \eta, \\ a_2 & \text{if } \eta \le x < 1, \end{cases}$$

the solution is

$$u(x) = \begin{cases} \left(\frac{1}{\eta+(1-\eta)a_1/a_2}\right) x & \text{if } 0 < x < \eta, \\ \left(\frac{1}{\eta a_2/a_1+(1-\eta)}\right) x + \left(1 - \frac{1}{\eta a_2/a_1+(1-\eta)}\right) & \text{if } \eta \leq x < 1. \end{cases} \tag{4.28}$$

This solution is linear in each section; at $x = \eta$, it satisfies the continuity and the jump conditions

$$u|_{x=\eta^-} = u|_{x=\eta^+}, \qquad a_1 \frac{du}{dx}\bigg|_{x=\eta^-} = a_2 \frac{du}{dx}\bigg|_{x=\eta^+}.$$

4.3.4 Numerical Results

When the singularity is severe, the standard FEM converges very slowly as described later in the section. This makes it hard to obtain the highly-resolved "exact" solution to which we can compare other computations. However, there is a way to measure the convergence rate without this resolved solution by comparing the errors on a sequence of grids, as long as there are no pointwise oscillations as the grid is refined. If we denote the solutions at the grids of dimensions $n, 2n$, and $4n$ by u_n, u_{2n}, and u_{4n}, respectively, we have

$$\|u_n - u_{2n}\| \leq Ch^\beta \left(\frac{1}{2^\beta} - 1\right), \tag{4.29}$$

$$\|u_{2n} - u_{4n}\| \leq C\left(\frac{h}{2}\right)^\beta \left(\frac{1}{2^\beta} - 1\right), \tag{4.30}$$

based on

$$\|u_n - u_{exact}\| \leq Ch^\beta.$$

Solving for the convergence rate β, we get

$$\beta = (\log 2)^{-1} \log\left(\frac{\|u_n - u_{2n}\|}{\|u_{2n} - u_{4n}\|}\right). \tag{4.31}$$

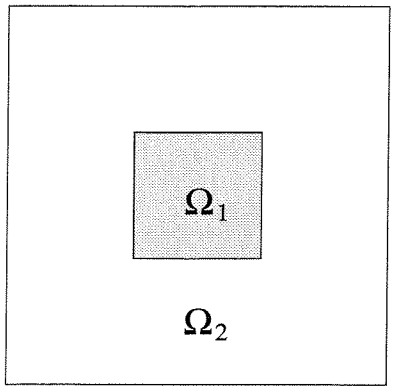

Figure 4.12: An interface with sharp corners; different coefficients in regions Ω_1 and Ω_2.

Corner Problem

We first consider singularities that arise from sharp corners in the interface. One example is shown in Figure 4.12. Suppose we have the coefficient

$$a(x,y) = \begin{cases} 1 & \text{if } x \in \Omega_1, \\ \alpha & \text{if } x \in \Omega_2. \end{cases}$$

For this problem, the standard FEM converges with $\mathcal{O}(h)$, when the grids are lined up with the interface [38]. This is easy to do if the interface is simple; if the interface is complicated, fitting the mesh around the interface is not a trivial process. If the singularity happens to be inside an element, the standard FEM results in a large error.

With the infinite elements at the four corners of the square, the convergence rate is still first order in term of h. However, the error is much smaller because the singularities are now well-resolved. We see in Table 4.3 that the l^2 error from the infinite element method is about 10 times smaller than that produced by the standard FEM, even with moderate $m = 16$ for the infinite element. A major of advantage of our new method is also the following: since the main source of the error is the singularity at the corners, we can keep the uniform coarse grid in the whole domain and simply refine the few infinite element bases locally. Unlike the standard adaptive mesh methods in which new refinement points are coupled globally to all the others, the refinement inside an infinite element is decoupled from the rest of the domain. The size of the global stiffness matrix does not change. We

n	m	l^2 error	ratio	l^∞ error	ratio
FEM 8	32	.07619		.17850	
IEM 8	16	.00733		.02946	
IEM 8	32	.00430	.7695	.01515	0.9594
IEM 8	64	.00299	.5242	.00768	0.9801

Table 4.3: Improvements in error with the infinite element bases at the corners.

α_1	α_2	l^2 error	ratio	l^∞ error	ratio
1	.001	0.0042230		0.015367	
1	.0001	0.0041871	0.9915	0.015252	0.9925
1	.00001	0.0039059	0.9328	0.013827	0.9066
1	.000001	0.0038436	0.9840	0.012988	0.9393

Table 4.4: Sensitivity of the infinite element basis method for the corner problem ($n = 8, m = 32$).

can see in Table 4.3 that the reduction in error with this refinement is nearly first order in h_m in the l^∞ norm, where $h_m = h/m$ is the mesh size along the boundary of the element. With this strategy, we can obtain very small errors.

We note that increasing m (an element has m points along each edge) here is different from the refinement of multiscale basis of the previous chapters. In the infinite element case, the number of unknowns is only doubled when m is doubled, since only the boundary points are involved in the computation. In the multiscale and most other methods, the number of unknowns is squared when m is doubled. Therefore, we can take m to be large without having to worry about the memory size.

Because the infinite element grid can resolve severe singularities, the error should also be insensitive to high contrast. Indeed, as we vary the contrast of the coefficients, the error stays basically the same up to $\alpha = 10^{-6}$ or so, as shown in Table 4.4.

Checkerboard Problem

Now we turn to the more difficult case of the checkerboard problem, shown in Figure 4.8. The leading order behavior at the singularity is $r^\gamma with \gamma = \sqrt{\alpha}$, as described earlier.

α	$\gamma = \sqrt{\alpha}$	$n = 64$	$n = 128$	$n = 256$	$n = 512$
.25	.5	1.023	1.035	1.038	1.040
.0625	.25	.619	.621	.613	.613
.015625	.125	.541	.462	.397	.366

Table 4.5: Convergence rates for standard FEM with a single juncture at $(.5, .5)$.

For a fixed point in the domain, the convergence of the standard finite element or finite difference method is predicted to be $\mathcal{O}(h^{2\gamma})$ [75]. This means that as contrast in the conductivities increases, the method is arbitrarily slow. This is already with the assumption that the gridlines are lined up with the discontinuities and the four-corner is on a grid point.

The numerical results shown in Table 4.5 of the convergence rates for the standard finite element method corroborate the prediction. At each n and α, we have estimated the convergence rate by comparing the l^2 error on successive grids. We see that as n gets large, the order settles down. By the last column, the rate is close to being proportional to 2γ, as the convergence rate is also halved when γ is halved. Thus the method slows down dramatically with the contrast.

In Figure 4.13 and Figure 4.14, we plot examples of the infinite element basis for the checkerboard case. These elements have 1 at the upper left corner and 0 at the remaining corners. The values on the boundary are obtained by solving the one-dimensional equation (4.28). The contrast is 10 for the first case and 10^5 for the second. We can see the difference in the degree of the singular behavior.

A good way to verify the performance of the infinite element bases is by considering the problem with multiple junctures, as in Figure 4.15. The coefficient alternates between α_1 and α_2 in the shaded and unshaded regions. For convenience, we again fix $\alpha_1 = 1$ and let α_2 be the parameter α. We lay down a grid in such a way that each juncture is placed in the middle of an element. This is not necessary for the method, but the implementation is a bit easier. The singular behavior is usually confined to the immediate vicinity of the singular point. As long as the element covers this area, the location of the singularity within the element is irrelevant.

In Table 4.5, we saw that the standard FEM performs very poorly with respect to the

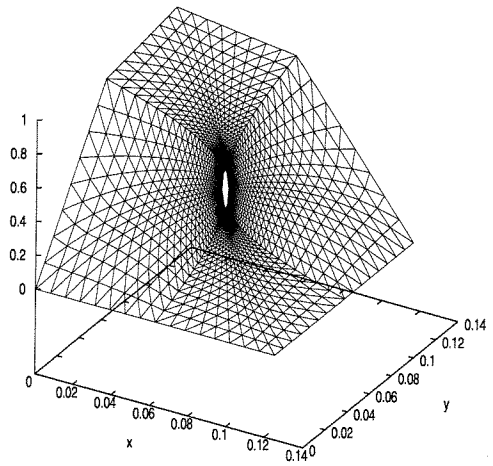

Figure 4.13: An infinite element basis for the checkerboard pattern, with conductivities $\alpha_1 = 1$ and $\alpha_2 = .1$ ($m = 16$); first 30 layers are plotted.

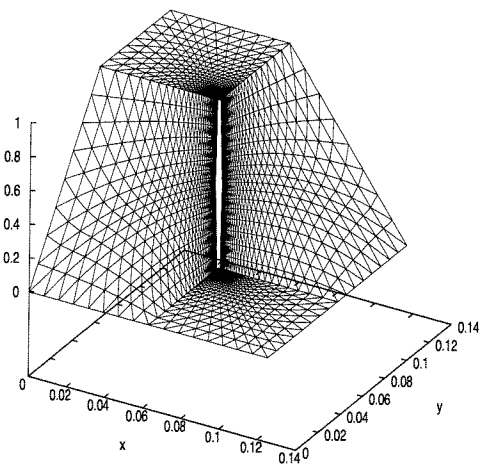

Figure 4.14: An infinite element for the checkerboard pattern, with conductivities $\alpha_1 = 1$ and $\alpha_2 = .00001$ ($m = 16$); first 30 layers are plotted.

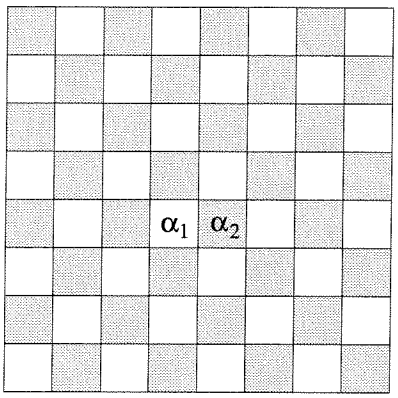

Figure 4.15: A checkerboard pattern with conductivities α_1 and α_2 alternating.

α	l^2 error	ratio	l^∞ error	ratio
e-1	0.0006797		0.0010896	
e-2	0.0048499	7.14	0.0078976	7.24
e-3	0.0144204	2.97	0.0239461	3.03
e-4	0.0172374	1.20	0.0300191	1.25
e-5	0.0166001	0.96	0.0309556	1.03
e-6	0.0166416	1.00	0.0311084	1.00
e-7	0.0177008	1.06	0.0311447	1.00

Table 4.6: Sensitivity to the contrast with the infinite element basis ($n = 16, m = 16$).

n	m	l^2 error	rate	l^∞ error	rate
$\alpha = 1/4$					
16	8	0.00067777		0.00107151	
	16	0.00032362	1.0665	0.00051008	1.0708
	32	0.00015251	1.0854	0.00024004	1.0874
	64	0.00007301	1.0626	0.00011484	1.0636
	128	0.00003556	1.0380	0.00005591	1.0385
$\alpha = 1/256$					
16	8	0.01290978		0.02175406	
	16	0.00888366	0.5392	0.01456563	0.5787
	32	0.00398064	1.1582	0.00644239	1.1769
	64	0.00129888	1.6157	0.00209229	1.6225
	128	0.00037668	1.7859	0.00060590	1.7879

Table 4.7: Error for multiple juncture problem as the infinite element basis is refined ($\alpha = 1/4, 1/256$).

increase in contrast. The convergence is arbitrarily slow as $\alpha \to 0$. In Table 4.6, we show the behavior of the error for the infinite element basis method as the contrast is increased. We see that the error increases initially, but grows less sensitive as the contrast is increased. Past $\alpha = .0001$ or so, the error essentially remains the same. We can explain this in terms of the area affected by the singularity. As the contrast grows larger, the singular behavior gets more severe, but it also gets more localized. We can see this from the analytical estimate (4.12) and from Figure 4.13 and Figure 4.14. Therefore, once the singular behavior is sufficiently localized due to high enough contrast, the infinite element captures this with very little error. The insensitivity to contrast due to the infinite element mesh is the great advantage of this method.

As described for the corner problem, we can again refine the local mesh inside the element, rather than the global mesh, in order to gain a substantial error reduction. In Table 4.7, we show how the overall error changes when the basis functions are refined. We see that for the relatively small contrast of $\alpha = 1/4$, the convergence rate with respect to the local grid size h_m is close to 1; for $\alpha = 1/256$, the rate actually increases, and approaches to 2. This shows that the overall error is indeed coming from the lack of refinement at the singularity and that the infinite element method can deal with the severe singularities very well.

We already mentioned that, in general, it is hard to obtain a well-resolved solution when

	n	m	l^2 error	% error	l^∞ error	% error
linear	64	16	0.950445	93%	2.183628	93%
IEM	64	8	0.146707	14%	0.379787	16%
	64	16	0.064785	6.4%	0.189092	8.1%
	64	32	0.032905	3.2%	0.106049	4.5%
	64	64	0.023659	2.3%	0.067518	2.9%

Table 4.8: Comparisons to the homogenized solution obtained using $a^* = \sqrt{\alpha_1 \alpha_2}$.

there is a strong singularity. However, for the checkerboard problem, we know that the homogenized coefficient is $\sqrt{\alpha_1 \alpha_2}$, where α_1 and α_2 are the alternating coefficients [52]. The final example uses this result to verify the performance of the infinite element method. We pick some boundary condition and forcing, e.g., $f(x, y) = 1$ in the domain and $u(x, y) = \sin(1.5\pi x)\sin(\pi(y + .5))$ on the boundary.

In Table 4.8, we show how different methods perform. We let $n = 64$ and $\alpha_1 = 1, \alpha_2 = .001$ in this case. We see that the linear method completely fails to capture the correct behavior, and the error is $\mathcal{O}(1)$. On the other hand, using the infinite elements gives the correct averaging, and the error is very small, down to 2.3% when $m = 64$. As before, the solution improves as the mesh inside the element is refined. As α_2 decreases, the size of the solution grows, but the relative error stays nearly the same. The relative errors are $8.9\%, 8.1\%, 7.3\%$ for $\alpha_2 = .01, .001, .0001$, respectively. For these computations, the boundary conditions used for the infinite element bases are again the piecewise constant solutions (4.28); using linear functions gives much larger errors.

4.4 Summary

In the earlier chapters, we developed finite element methods with basis functions that capture the solution behavior accurately. In this chapter, the same approach has been extended to elliptic problems with singular behavior. We have considered a high contrast but smooth coefficient for which asymptotic results can be used to construct the basis functions. For the discontinuous coefficients, we have used the infinite element method for the elements that contain the singularity. In particular, we have considered the corner

and checkerboard singularities. Because most of the error comes from the neighborhood of singularities, we found that the local refinement of the infinite element results in the reduction of the global error and that this can be done efficiently involving only the points on the element boundary. Finally, we have found that the infinite element approach gives excellent results, without regard to the contrast.

Chapter 5 Conclusions

Efficient numerical methods have been developed for dealing with the boundary and internal layer structure in the convection-diffusion equation and for the singular behavior in the elliptic equation. Simple extensions of successful one-dimensional approaches are not capable of capturing complicated two-dimensional effects, and general adaptive mesh algorithms lead to large, often ill-conditioned matrix problems. An effective alternative approach we propose in the present work is to construct the basis functions in the finite element setting that contain the correct local behavior.

We find that there are two cases in general. In Chapter 2, we deal with the case when the characteristic length scale of the problem is small enough to be contained in an element basis. Then the multiscale finite element method works well, as the solution of the homogeneous equation solved locally reflects the global behavior correctly. This is the case in the periodic cellular flow we consider. We analyze this multiscale method using multiple scale expansions and homogenization estimates. To obtain the correct estimate of the convergence rate, we need to consider the subtle cancellation of errors in the discrete equation. We find that for the cellular flow, the leading order term of the relative error is $\mathcal{O}(\delta/h)$, where δ is the spatial oscillation and h the mesh spacing. Thus the method converges as $\delta \to 0$ for a fixed h, regardless of the small diffusivity ϵ. We then try to improve the multiscale method further by constructing asymptotic basis functions, since solving them numerically may become expensive in the limit $\epsilon \to 0$. This involves a transformation to the streamline coordinates, on which the resulting Wiener-Hopf problem has already been worked out. Although we can construct such bases successfully for the cellular flow, numerical methods will have to be employed for obtaining the basis in general.

In Chapter 3, we perturb the cellular structure randomly to obtain a random flow. The resulting characteristic length scale is much larger as estimated by percolation theory. The layer structure is nonlocal in this case and the multiscale basis function cannot produce the

necessary basis functions by sampling a small region. However, we find that a variational principle has been derived for this situation and can provide the necessary information if employed properly. The variational principle is still nonlocal, but the correct scaling of the layer that minimizes the energy can be found efficiently using some test functions that capture the exponential layer structure. Using this information, exponential basis functions can be constructed and added to the finite element space. This destroys the banded structure of the stiffness matrix for bilinear finite elements, but the matrix problem can still be solved efficiently. When these extra functions are incorporated correctly, there is a substantial improvement in the numerical solution.

The same approach of adaptive basis functions provides a natural setting for dealing with the singular behavior of an elliptic problem when the coefficient has a high contrast. This case is studied in Chapter 4. When the coefficient has a saddle point behavior resulting in a narrow channel of high concentration, we can capture its behavior correctly using an asymptotic basis function based on an asymptotic result. For the discontinuous interface problems with high contrast, the basis functions can be constructed using the infinite element method. Based on the self-similarity of a specialized grid, it provides an accurate way of capturing the behavior near a singular point. We apply this method to interfaces with corners and the checkerboard problem and see that the method is efficient and robust.

Throughout this work, the main idea has been that when the finite element space is enriched with the basis functions that capture the correct local behavior, we can use a relatively coarse grid and yet still capture various singular behavior in a robust manner. This approach is particularly well-suited for periodic velocity fields such as the cellular flow and interface problems with sharp singularities, for which conventional methods require a prohibitively large number of grid points. Other problems, particularly those involving the random flows of Chapter 3, have additional difficulties such as nonlocality and nonuniformity of layers. We can overcome these problems to some extent through a variational principle and analytical results. Further research would involve obtaining better predictions regarding the location of the layers and dealing with layers that evolve in time, as well as applying the methods developed in this work to problems in other areas.

Appendix A Proof of Lemma 2.1

For the one-dimensional convection-diffusion equation

$$-u'' + a(x)\, u' = f, \quad x \in [0,1], \tag{A.1}$$

with boundary conditions $u(0) = u(1) = 0$, we substitute $u(x) = \sum_j u_j \phi_j(x)$ and $v = \psi_i(x)$ in the weak form

$$(u', v') + (au', v) = (f, v), \quad \forall v \in H_0^1 \tag{A.2}$$

to get

$$\sum_j \left[(\phi_j', \psi_i') + (a\phi_j', \psi_i)\right] u_j = (f, \psi_i). \tag{A.3}$$

With integration by parts,

$$(\phi_j', \psi_i') + (a\phi_j', \psi_i) = (\phi_j, \psi_i')|_{x_{j-1}}^{x_{j+1}} - (\phi_j, \psi_i'') - (\phi_j, (a\psi_i)') \tag{A.4}$$

$$= (\phi_j, \psi_i')|_{x_{j-1}}^{x_{j+1}}, \tag{A.5}$$

if the trial functions are the "hat" functions as defined in (2.10), or some other continuous functions with the correct boundary conditions, and the test functions are the solutions of the adjoint equation

$$-\psi_i'' - (a(x)\psi_i)' = 0. \tag{A.6}$$

We could integrate in the other direction in (A.4) and then pick the trial functions that satisfy $-\phi_j'' + (a(x)\phi_j)' = 0$, but this does not give the desired result in the end when forcing is present.

Defining $b(x) \equiv \exp(\int^x a(s)\, ds)$, we find the exact expressions for this test function in (A.6)

to be

$$\psi(x) = \begin{cases} \left(b(x_i)\int_{x_{i-1}}^{x} b(s)\,ds\right) \Big/ \left(b(x)\int_{x_{i-1}}^{x_i} b(s)\,ds\right), & x \in [x_{i-1}, x_i], \\ \left(b(x_i)\int_{x_i}^{x_{i+1}} b(s)\,ds\right) \Big/ \left(b(x)\int_{x_i}^{x_{i+1}} b(s)\,ds\right), & x \in [x_i, x_{i+1}]. \end{cases} \tag{A.7}$$

Now we can compute the inner product of (A.5) exactly and find that the discretized equations (A.3) are now

$$-\left(\frac{1}{\int_{x_{i-1}}^{x_i} b}\right) u_{i-1} + \left(\frac{1}{\int_{x_{i-1}}^{x_i} b} + \frac{1}{\int_{x_i}^{x_{i+1}} b}\right) u_i - \left(\frac{1}{\int_{x_i}^{x_{i+1}} b}\right) u_{i+1} = \frac{1}{b(x_i)} \int_{x_{i-1}}^{x_{i+1}} f\psi_i\,ds. \tag{A.8}$$

We will write $\int_{x_{i=1}}^{x_i} b(s)\,ds$ as $\int_{x_{i=1}}^{x_i} b$, etc., in the following for convenience.

We now expand the right-hand side using ψ_i to get

$$\frac{1}{b(x_i)} \int_{x_{i-1}}^{i+1} f\psi_i\,ds = \frac{1}{b(x_i)} \left(\int_{x_{i-1}}^{x_i} f\psi_i\,ds + \int_{x_i}^{x_{i+1}} f\psi_i\,ds \right) \tag{A.9}$$

$$= \frac{\int_{x_{i-1}}^{x_i} f(s) b^{-1}(s) \left(\int_{x_{i-1}}^{s} b\right) ds}{\int_{x_{i-1}}^{x_i} b} + \frac{\int_{x_i}^{x_{i+1}} f(s) b^{-1}(s) \left(\int_{x_i}^{s} b\right) ds}{\int_{x_i}^{x_{i+1}} b}. \tag{A.10}$$

If we define

$$F(x) \equiv \int_0^x f b^{-1} \quad \text{and} \quad G \equiv \int_0^x F b, \tag{A.11}$$

and integrate by parts, we have

$$\frac{1}{b(x_i)} \int_{x_{i-1}}^{i+1} f\psi_i\,ds = F(x_i) - \frac{\int_{x_{i-1}}^{x_i} Fb}{\int_{x_{i-1}}^{x_i} b} - F(x_i) + \frac{\int_{x_i}^{x_{i+1}} Fb}{\int_{x_i}^{x_{i+1}} b} \tag{A.12}$$

$$= \frac{1}{\int_{x_{i-1}}^{x_i} b}(G(x_{i-1}) - G(x_i)) + \frac{1}{\int_{x_i}^{x_{i+1}} b}(G(x_{i+1}) - G(x_i)) \tag{A.13}$$

$$= \frac{1}{\int_{x_{i-1}}^{x_i} b} G(x_{i-1}) + \left(\frac{1}{\int_{x_{i-1}}^{x_i} b} + \frac{1}{\int_{x_i}^{x_{i+1}} b}\right) G(x_i) - \frac{1}{\int_{x_i}^{x_{i+1}} b} G(x_{i+1}). \tag{A.14}$$

Now we combine this with (A.8) to obtain

$$\frac{1}{\int_{x_{i-1}}^{x_i} b}\left(u(x_{i-1})+G(x_{i-1})\right)+\left(\frac{1}{\int_{x_{i-1}}^{x_i} b}+\frac{1}{\int_{x_i}^{x_{i+1}} b}\right)\left(u(x_i)+G(x_i)\right) - \frac{1}{\int_{x_i}^{x_{i+1}} b}\left(u(x_{i+1})+G(x_{i+1})\right)=0. \quad \text{(A.15)}$$

Defining $E_j = u(x_j) + G(x_j)$, this reduces to

$$-\frac{1}{\int_{x_{i-1}}^{x_i} b}\left(E_i - E_{i-1}\right)+\frac{1}{\int_{x_i}^{x_{i+1}} b}\left(E_{i+1}-E_i\right)=0. \quad \text{(A.16)}$$

With the boundary condition $E_0 = 0$, we find that $E_i = c\int_{x_0}^{x_i} b$. With $E_N = u(x_N) + G(x_N) = \int_0^1 Fb$, we have $c = \int_0^1 Fb / \int_0^1 b$. Therefore, the values at the nodal points are

$$u(x_i) = E_i - G(x_i) \quad \text{(A.17)}$$

$$= -\int_0^{x_i} Fb + \frac{\int_0^1 Fb}{\int_0^1 b}\int_0^{x_i} b. \quad \text{(A.18)}$$

Meanwhile, the exact solution can be derived for the original equation (A.1). Using the integrating factor,

$$(b^{-1}u')' = -fb^{-1}, \quad \text{(A.19)}$$

$$u' = -bF + c_0 b, \quad \text{(A.20)}$$

and finally

$$u(x) = -\int_0^x bF + c_0\int_0^x b + c_1. \quad \text{(A.21)}$$

With boundary conditions $u(0) = u(1) = 1$,

$$c_0 = \frac{\int_0^x Fb}{\int_0^x b}, \quad \text{(A.22)}$$

$$c_1 = 0. \quad \text{(A.23)}$$

We see that the exact solution (A.21) at the nodal points then match the values obtained from the finite element method (A.18).

Bibliography

[1] R. E. Alcouffe, A. Brandt, J. E. Dendy, and J.W. Painter. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM Journal on Scientific and Statistical Computing*, 2(4):430–454, 1981.

[2] D. Allen and R. Southwell. Relaxation methods applied to determining the motion in two dimensions of a viscous fluid past a fixed cylinder. *Quarterly Journal of Mechanics and Applied Mathematics*, 8:129–145, 1955.

[3] A. P. Anufriyev and V. M. Fishman. Magnetic field structure in the two-dimensional motion of a conducting field. *Geomagnetism and Aeronomy*, 22(2):245–248, 1982.

[4] C. Apelian, R. L. Holmes, and M. Avellaneda. A turbulent transport model: Streamline results for a class of random velocity fields in the plane. *Communications on Pure and Applied Mathematics*, 50(11):1053–1088, 1997.

[5] M. Avellaneda. Enhanced diffusivity and intercell transition layers in 2-d models of passive advection. *Journal of Mathematical Physics*, 32(11):3209–3212, 1991.

[6] M. Avellaneda and A. J. Majda. Stieltjes integral-representation and effective diffusivity bounds for turbulent transport. *Physical Review Letters*, 62(7):753–755, 1989.

[7] M. Avellaneda and A. J. Majda. An integral-representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Communications in Mathematical Physics*, 138(2):339–391, 1991.

[8] M. Avellaneda, S. Torquato, and I. C. Kim. Diffusion and geometric effects in passive advection by random arrays of vortices. *Physics of Fluids*, 3(8):1880–1891, 1991.

[9] I. Babuška and A. K. Aziz. Finite element method for non-smooth domains and coefficients. In A. K. Aziz, editor, *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, pages 265–283. Academic Press, New York, 1972.

[10] I. Babuška and J. E. Osborn. Finite element methods for the solution of problems with rough input data. In P. Grisvard, W. Wendland, and J. R. Whiteman, editors, *Singularities and Constructive Methods for Their Treatment : Proceedings of the Conference Held in Oberwolfach, West Germany, November 20- 26, 1983*, pages 1–18. Springer-Verlag, Berlin, 1985.

[11] A. M. Balk and R. M. McLaughlin. Passive scalar in a random wave field: the weak turbulence approach. *Physics Letters A*, 256(4):299–306, 1999.

[12] R. E. Bank. *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations.* SIAM Books, Philadelphia, 1990.

[13] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition.* SIAM, Philadelphia, PA, 1994.

[14] G. K. Batchelor and R. W. O'Brien. Thermal or electrical conduction through a granular material. *Proc. R. Soc. London A*, 355:313–333, 1977.

[15] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers.* McGraw-Hill, New York, 1978.

[16] A. Bensoussan, J.–L. Lions, and G. C. Papanicolaou. *Asymptotic Analysis for Periodic Structures.* North Holland, Amsterdam, 1978.

[17] L. Borcea and G. C. Papanicolaou. A hybrid numerical method for high contrast conductivity problems. *Journal of Computational and Applied Mathematics*, 87(1):61–77, 1997.

[18] L. Borcea and G. C. Papanicolaou. Network approximation for transport properties of high contrast materials. *SIAM Journal of Applied Mathematics*, 58(2):501–539, 1998.

[19] J.–P. Bouchaud, A. Georges, J. Koplik, A. Provata, and S. Redner. Superdiffusion in random velocity fields. *Physical Review Letters*, 64(21):2503–2506, 1990.

[20] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, New York, 1994.

[21] C. J. Budd, G. P. Koomullil, and A. M. Stuart. On the solution of convection-diffusion boundary value problems using equidistributed grids. *SIAM Journal on Scientific Computing*, 20(2):591–618, 1999.

[22] G. F. Carrier, M. Krook, and C. E. Pearson. *Functions of a Complex Variable.* Hod Books, Ithaca, New York, 1983.

[23] Z. Chen and J. Zou. Finite element methods and their convergence for elliptic and parabolic interface problems. *Numerische Mathematik*, 79:175–202, 1998.

[24] S. Childress. Alpha-effect in flux ropes and sheets. *Physics of the Earth and Planetary Interiors*, 20:172–180, 1979.

[25] S. Childress and A. M. Soward. Scalar transport and alpha-effect for a family of cats-eye flows. *Journal of Fluid Mechanics*, 205:99–133, 1989.

[26] P. G. Ciarlet and J.–L. Lions. *Handbook of Numerical Analysis, Vol. II.* Elsevier, Amsterdam, 1990.

[27] G. Cottafava and G. Le Moli. Automatic contour map. *Communications of the ACM*, 12(7):386–391, 1969.

[28] P. M. De Zeeuw. Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *Journal of Computational and Applied Mathematics*, 33(1):1–27, 1990.

[29] J. Douglas, Jr. and T. F. Russell. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM Journal of Numerical Analysis*, 19(5):871–885, 1982.

[30] Ya. E. Efendiev, T. Y. Hou, and X. H. Wu. Convergence of a nonconforming multiscale finite element method. *to appear in SIAM Journal of Numerical Analysis.*

[31] A. Fannjiang and G. C. Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM Journal on Applied Mathematics*, 54(2):333–408, 1994.

[32] A. Fannjiang and G. C. Papanicolaou. Diffusion in turbulence. *Prob. Theory and Rel. Fields*, 105:279–334, 1996.

[33] A. Fannjiang and G. C. Papanicolaou. Convection-enhanced diffusion for random flows. *Journal of Statistical Physics*, 88(5–6):1033–1076, 1997.

[34] I. M. Gelfand and S. V. Fomin. *Calculus of Variations.* Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

[35] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag, Berlin, 1983.

[36] G. H. Golub, D. J. Silvester, and A. J. Wathen. Diagonal dominance and positive definiteness of upwind approximations for advection diffusion problems. In D. F. Griffiths and G. A. Watson, editors, *Numerical Analysis: A.R. Mitchell 75th Birthday Volume*, pages 125–131. World Scientific, Singapore, 1996.

[37] A. V. Gruzinov, M. B. Isichenko, and Ya. L. Kalda. Two-dimensional turbulent diffusion. *Sov. Phys. JETP*, 70(2):263–269, 1990.

[38] H. Han. The numerical-solutions of interface problems by infinite element method. *Numerische Mathematik*, 39(1):39–50, 1982.

[39] A. F. Hegarty, E. O'Riordan, and M. Stynes. A comparison of uniformly convergent difference-schemes for 2-dimensional convection diffusion-problems. *Journal of Computational Physics*, 105(1):24–32, 1993.

[40] E. J. Hinch. *Perturbation Methods.* Cambridge University Press, Cambridge, 1991.

[41] M. H. Holmes. *Introduction to Perturbation Methods.* Springer-Verlag, New York, 1995.

[42] M. Holzer and E. D. Siggia. Turbulent mixing of a passive scalar. *Physics of Fluids*, 6(5):1820–1837, 1994.

[43] T. Y. Hou and X. H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134(1):169–189, 1997.

[44] T. Y. Hou, X. H. Wu, and Z. Q. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation*, 68(227):913–943, 1999.

[45] M. B. Isichenko. Percolation, statistical topography, and transport in random-media. *Reviews of Modern Physics*, 64(4):961–1043, 1992.

[46] M. B. Isichenko and Ya. L. Kalda. Statistical topography 1: Fractal dimension of coastlines and number-area rule for islands. *Journal of Nonlinear Science*, 1:256–277, 1991.

[47] M. B. Isichenko and Ya. L. Kalda. Statistical topography 2: Two-dimensional transport of a passive scalar. *Journal of Nonlinear Science*, 1:375–396, 1991.

[48] M. B. Isichenko, Ya. L. Kalda, E. B. Tatarinova, O. V. Tel'kovskaya, and V. V. Yan'kov. Diffusion in a medium with vortex flow. *Sov. Physics JETP*, 69(3):517–524, 1989.

[49] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, Berlin, 1991.

[50] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, Great Britain, 1987.

[51] H. B. Keller. *Lectures on Numerical Methods in Bifurcation Problems.* Springer-Verlag, Berlin, for the Tata Institute of Fundamental Research, 1987.

[52] J. B. Keller. Effective conductivity of periodic composites composed of two very unequal conductors. *Journal of Math. Phys.*, 28:2516–2520, 1987.

[53] R. B. Kellogg and A. Tsan. Analysis of some difference approximations for a singular perturbation problem without turning points. *Mathematics of Computation*, 32:1025–1039, 1978.

[54] J. Kevorkian and J. D. Cole. *Multiple Scale and Singular Perturbation Methods.* Springer, New York, 1996.

[55] S. M. Kozlov. Geometric aspects of averaging. *Russian Math. Surveys*, 44(2):91–144, 1989.

[56] R. H. Kraichnan. Diffusion by a random velocity field. *The Physics of Fluids*, 13(1):22–31, 1970.

[57] R. J. Leveque. *Numerical Methods for Conservation Laws.* Birkhäuse, Basel, 1992.

[58] R. J. Leveque and Z. L. Li. The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM Journal on Numerical Analysis*, 31(4):1019–1044, 1994.

[59] G. Lube. An asymptotically fitted finite element method for convection dominated convection-diffusion-reaction problems. *Zeitschrift fur Angewandte Mathematik und Mechanik*, 72(3):189–200, 1992.

[60] D. W. McLaughlin, G. C. Papanicolaou, and O. R. Pironneau. Convection of microstructure and related problems. *SIAM Journal of Applied Mathematics*, 45:780–797, 1985.

[61] C. C. Mei, J.–L. Auriault, and C. Ng. Some applications of the homogenization theory. *Advances in Applied Mechanics*, 12:277–349, 1996.

[62] J. J. H. Miller, editor. *Boundary and Interior Layers: Computational and Asymptotic Methods: proceedings of the BAIL conference held at Trinity College, Dublin, June 3–6, 1980.* Boole Press, Dublin, 1980.

[63] J. J. H. Miller, E. O'Riordan, and G. I. Shishkin. *Fitted Numerical Methods for Singular Perturbation Problems.* World Scientific, Singapore, 1996.

[64] G. W. Milton. Classical hall-effect in two-dimensional composites - a characterization of the set of realizable effective conductivity tensors. *Physical Review B-Condensed Matter*, 38(16):11296–11303, 1988.

[65] H. K. Moffatt. *Magnetic Field Generation in Electrically Conducting Fluids.* Cambridge University Press, Cambridge, 1978.

[66] H. Oh and I. Babuška. The p-version of the finite element method for the elliptic boundary value problems with interfaces. *Computer Methods in Applied Mechanics and Engineering*, 97(2):211–231, 1992.

[67] R. E. O'Malley, Jr. *Introduction to Singular Perturbations.* Academic Press, New York, 1974.

[68] E. O'Riordan and M. Stynes. A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in two dimensions. *Mathematics of Computation*, 57:47–62, 1991.

[69] S. J. Polak, C. den Heijer, and W. H. A. Schilders. Semiconductor device modelling from the numerical point of view. *International Journal of Numerical Methods in Engineering*, 24:763–838, 1987.

[70] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: the art of scientific computing.* Cambridge University Press, Cambridge, England, 1992.

[71] J. N. Reddy. *An Introduction to the Finite Element Method.* McGraw-Hill, New York, 1984.

[72] H. G. Roos, D. Adam, and A. Felgenhauer. A novel nonconforming uniformly convergent finite element method in two dimensions. *Journal of Mathematical Analysis and Applications*, 201(3):715–755, 1996.

[73] H.-G. Roos, M. Stynes, and L. Tobiska. *Numerical Methods for Singularly Perturbed Differential Equations.* Springer-Verlag, Berlin, 1991.

[74] M. N. Rosenbluth, H. L. Berk, I. Doxas, and W. Horton. Effective diffusion in laminar convective flows. *Physics of Fluids*, 80:2636–2647, 1987.

[75] U. Rüde and C. Zenger. On the treatment of singularities in the multigrid method. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods II*, pages 261–271. Springer-Verlag, Berlin, 1986.

[76] R. Sacco and M. Stynes. Finite element methods for convection-diffusion problems using exponential splines on triangles. *Computers and Mathematics with Applications*, 35(3):35–45, 1998.

[77] E. Sanchez-Palencia. *Non-homogenous Media and Vibration Theory.* Springer-Verlag, New York, 1980.

[78] A. M. Soward. Fast dynamo action in a steady flow. *Journal of Fluid Mechanics*, 180:267–295, 1987.

[79] G. W. Strang and G. J. Fix. *An Analysis of the Finite Element Method.* Prentice–Hall, Englewood Cliffs, N.J., 1973.

[80] J. C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Chapman and Hall, New York, 1989.

[81] M. Stynes and E. O'Riordan. An analysis of a singularly perturbed two-point bouundary value problems using only finite element techniques. *Mathematics of Computation*, 56:663–675, 1991.

[82] L. Ying. *Infinite Element Methods.* Peking University Press, Beijing, 1995.

[83] R. M. Ziff. Test of scaling exponents for percolation-cluster perimeter. *Physical Review Letters*, 56(6):545–548, 1986.