THE K-GRID FOURIER ANALYSIS OF MULTIGRID-TYPE ITERATIVE METHODS

by

Naomi H. Decker

Computer Sciences Technical Report #703

July 1987

THE K-GRID FOURIER ANALYSIS OF MULTIGRID-TYPE ITERATIVE METHODS[1]

by

Naomi H. Decker

Computer Sciences Technical Report #703

[1] Sponsored by the Air Force Office of Scientific Research under Contracts No. AFOSR-82-0275 and 86-0163.

ABSTRACT

Experiments indicate that a multigrid-type cycle can be used as an efficient preconditioner in the iterative solution of the discrete problem corresponding to a singularly perturbed elliptic boundary value problem. Motivated by a report of Goldstein, we explore the theoretical basis for the efficiency of such a preconditioner when applied to a model problem. The techniques developed are also used to analyze a multigrid V-cycle when used alone as a fast iterative solver.

1. Introduction

This work is motivated by a report of Charles Goldstein [7] in which the author discusses the task of numerically solving the following elliptic boundary value problem:

$$\begin{cases} -\varepsilon^2 \sum_{i=1}^{2} \frac{\partial}{\partial x_i}\left(a_i(x)\frac{\partial u(x)}{\partial x_i}\right) + \varepsilon \sum_{i=1}^{2} b_i(x)\frac{\partial u(x)}{\partial x_i} + a_0(x)u(x) = f(x) \text{ in } \Omega \subset \mathbb{R}^2 \\ \\ u(x) = g(x) \text{ on } \partial\Omega \end{cases} \tag{1.1}$$

where $x = (x_1, x_2) \in \Omega$, $0 < \varepsilon << 1$, the coefficients and data are sufficiently smooth, and $a_i(x) > c_0 > 0$ in Ω , $i = 0, 1, 2$.

The discrete problem arising from a typical discretization of (1.1) on a uniform grid of mesh size h, $h < \varepsilon$, is a large system of linear equations. For the solution of this system to approximate the solution of the boundary value problem (1.1) with a fixed accuracy, we must choose the mesh size small for small ε, specifically, it is sufficient to keep the ratio h/ε fixed [1], [11]. In doing so, we not only get a much larger system, but the resulting system is also more poorly conditioned.

With the goal of trying to solve this type of system, we use the conjugate gradient algorithm as our iterative solver. It is known (e.g., [2],[9]) that if we apply the method of conjugate gradients to the problem $Bv = F$ where B is symmetric, positive definite, then the number of iterations, N_B, required to solve the system to within a given relative error, $\|v - v^i\|/\|v - v^0\| < \eta$, is given by

$$N_B(\eta) \leq Cln(2/\eta)\sqrt{K(B)} \tag{1.2}$$

where $K(B) = \lambda_{\max}(B)/\lambda_{\min}(B)$, v^0 is the initial guess and v^i is the i th approximant to the solution, v. Our goal is to precondition the system so that the condition number, $K(B')$, of the new system, $B'v' = F'$, is much smaller than $K(B)$ and behaves nicely (bounded or slowly increasing) as ε and h decrease to zero.

It has been observed experimentally that a certain multigrid-type cycle is an inexpensive preconditioner for this system. The effectiveness of this preconditioner is quite sensitive to the choice of the number of grids, k, used in the multigrid process. Fourier analysis was used in [7] in an attempt to prove that a careful choice of the number of grids does guarantee a good preconditioner in the case where Ω is a rectangle. Although Fourier analysis is routinely used to study 2-grid multigrid cycles, the k-grid analysis, for $k > 2$, is quite unwieldly and is not usually attempted. The difficulty arises from the use of coarser grids on which certain modes "alias" (see [3]) or are "not visible" (see [12]). Unfortunately, this "aliasing" was ignored in [7]. The experimental evidence is so striking, however, that it seemed worth trying to complete the analysis.

We examine the effectiveness of the multigrid preconditioner by considering a special case of the boundary value problem (1.1) with $a_i(x) \equiv 1,\ i = 0,1,2$, $b_i(x) \equiv 0,\ i = 1,2$, $\Omega = (0,1)\times(0,1)$ and ε real and small. It is for this model operator, $A_L^\varepsilon = -\varepsilon^2\Delta + I$, that we prove our basic results. More general singularly perturbed problems such as variable coefficient and/or non–symmetric with positive definite symmetric part can be analyzed using the properties of the multigrid preconditioner acting on A_L^ε together with such ideas as spectral or norm equivalence, see [5] and [7].

Let $h = 2^{-n}$ for a positive integer, n. Discretizing this model problem on a uniform grid, $\Omega_h = \{(lh, mh) : l, m = 1, 2, \ldots, 2^n - 1\}$, with mesh size, h, using a standard 5-point discretization of the Laplacian (see Section 2.1), we obtain the linear system

$$A_h^\varepsilon u_h := (-\varepsilon^2 \Delta_h + I)u_h = f_h. \tag{1.3}$$

In Section 3.1 we define a symmetric linear operator, M_k, based on multigrid ideas, using $k-1$ auxiliary grids of larger mesh sizes, $2^p h$, for $p = 1, 2, \ldots, k-1$. In fact, the vector $M_k w_h$ is essentially one "partial" multigrid V-cycle applied as if to solve the problem:

$$A_h v_h = w_h, \tag{1.4}$$

starting with initial guess $= 0$, where A_h is the matrix resulting from the corresponding discretization of the Dirichlet boundary value problem for Poisson's equation. In order to obtain a symmetric operator, we take symmetric smooths. I.e., if r_p smooths are done on the p th grid in the fine to coarse part of the cycle, then r_p smooths must be done on the p th grid in the coarse to fine part. We take a fixed $r_p = r$ for all $p = 0, \ldots, k-1$. The adjective "partial" refers to the following property of this particular V–cycle: instead of solving for the coarse grid correction exactly on the coarsest grid, $2r$ iterations of the smoother are applied. We choose the smoother to be a damped Jacobi iteration with damping parameter, ω, where $0 < \omega < 1$. Taking $\omega = 1$ would correspond to an undamped Jacobi iteration, but we exclude this choice. The choice $\omega = .5$ corresponds to a Richardson iteration. Using M_k as a preconditioner for (1.3), we claim:

If the mesh size on the coarsest grid is choosen to be approximately equal to the singular perturbation parameter, ε, then the condition number of the preconditioned system is bounded independent of ε and h.

Defining $M_h^\varepsilon = M_k$, where k is chosen so that $h_1 \approx \varepsilon$, we justify this claim in 3 steps:

1. In Section 3.2 we reduce the problem to finding appropriate upper and lower bounds for the eigenvalues of $M_h^\varepsilon A_h^\varepsilon$. Let $q : \Omega_h \to \{1, 2, \ldots, (2^n - 1)^2\} : (i_1 h, i_2 h) \mapsto q_i, i = (i_1, i_2)$, be a given ordering of the $(2^n - 1)^2$ points of Ω_h, and let $\{\alpha_i\}$ be a (given) complete set of eigenvectors of A_h. Define a $(2^n - 1)^2 \times (2^n - 1)^2$ matrix, $\mathcal{M}$, by

$$(\mathcal{M})_{q_i, q_j} = \mu_{ij}$$

   where

$$\mu_{ij} := \langle M_h^\varepsilon A_h^\varepsilon \alpha_i, \alpha_j \rangle$$

   for each $i = (i_1, i_2)$, $j = (j_1, j_2)$ where $1 \le i_1, i_2, j_1, j_2 < 2^n$ and $\langle \cdot, \cdot \rangle$ is the discrete - L^2 inner product. Using this eigenfunction analysis (Fourier analysis), the problem reduces to finding bounds on the eigenvalues of $\mathcal{M}$. The off-diagonal elements of $\mathcal{M}$ represent the "aliasing".
2. In Section 3.3 we obtain a formula for a bound, $C^i_{h,k,r,\omega}$, such that, for every i,

$$\sum_{j \ne i} |\mu_{ij}| \le C^i_{h,k,r,\omega} |\mu_{ii}|.$$

   Therefore we have diagonal dominance of the matrix, $\mathcal{M}$, provided $\bar{C}_{h,k,r,\omega}$, where

$$\bar{C}_{h,k,r,\omega} := \sup_i C^i_{h,k,r,\omega},$$

can be shown to be less than one. The constant $\bar{C}_{h,k,r,\omega}$ is calculated for $r = 1,2,3,4$, $\omega = .5,.6,.7,.8,.9$, $h = 1/2, 1/4, 1/8, \ldots, 1/8192$ and all possible corresponding values of k. All computed values of $\bar{C}_{h,k,r,\omega}$ are less than one with the exception of the case where only one smoothing is used and $\omega < .7$.

3. In Section 3.5 we restate the bounds given in [7] on the diagonal entries of the matrix. These bounds are used, combined with the diagonal dominance, to show that:

$$c_1\varepsilon^2 \leq \lambda_{\min}(M_h^\varepsilon A_h^\varepsilon) \leq \lambda_{\max}(M_h^\varepsilon A_h^\varepsilon) \leq c_2\varepsilon^2,$$

for constants $c_1, c_2 > 0$. The diagonal dominance of $\mathcal{M}$ is needed only to guarantee the positivity of the lower bound.

In Section 4 we describe some experiments which illustrate the efficiency of using the optimal number of grids in the multigrid preconditioner. Experimental comparisons are made between three different solvers for the model problem. In a preconditioned conjugate gradient routine, two preconditioners are used, first the preconditioner analyzed in this paper, namely the preconditioner based on the Laplacian with smoothing on the coarsest grid, and secondly a preconditioner which is based on the model operator itself, solving on the coarse grid. The third solver used in the comparison is a symmetric multigrid V-cycle.

The techniques used in the analysis of "multigrid–as–a–preconditioner" can also be used to analyse "multigrid–as–a–solver". This analysis is simpler than the preconditioner analysis since we don't need diagonal dominance (and we don't have it), see Section 5. In Section 6 we show how the k-grid convergence bounds obtained in this way compare to the experimentally observed convergence rates and to V-cycle convergence bounds obtained by other methods.

2.1 Notation

Consider the two-dimensional Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega = (0,1)\times(0,1) \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{2.1}$$

where $\Delta = \sum_{j=1}^{2} \partial^2/\partial x_j^2$. We discretize this problem on a family of grids. Let $h = 2^{-n}$, as in Section 1. Choose a positive integer k, $k < n$. Define a coarse grid mesh size $h_1 = 2^{k-1}h$. In Ω we define k intermediate grids, Ω^p, $p = 1,2,...,k$ with mesh sizes $h_p = 2^{1-p}h_1$. Clearly $h = h_k$ and

$$\Omega^p = \{(x_l, y_m) = (lh_p, mh_p) : l, m = 1,2,...,N_p - 1\} \tag{2.2}$$

where $N_p = 1/h_p$ and $p = 1,2,...,k$.

We define the discrete operator, A_p, which is the negative of the discrete five point Laplacian, on the grid Ω^p, using the standard five-point discretization of the differential operator, $-\Delta$ (see e.g., [6]). Each A_p is a sparse $(N_p-1)^2 \times (N_p-1)^2$ matrix with a complete set of eigenvectors, $\alpha_i^{(p)}$, given by:

$$\alpha_i^{(p)}(m,n) = 2\sin(i_1\pi m h_p)\ \sin(i_2\pi n h_p) \qquad m,n = 1,...,N_p - 1\,. \tag{2.3}$$

where $i = (i_1, i_2)$, and $i_1, i_2 = 1,2,...,N_p - 1$. The corresponding eigenvalues are:

$$\nu_i^{(p)} = \frac{4 - 2\ \cos(i_1\pi h_p) - 2\ \cos(i_2\pi h_p)}{h_p^2}\,. \tag{2.4}$$

As usual, the multigrid operators we consider are constructed from smoothers, G_p, $p = 1,2,...,k$ and intergrid transfer operators, I_{p-1}^p and I_p^{p-1}, $p = 2,3,...,k$.

To simplify the analysis we choose $G_p(\cdot,\cdot)$ to be a damped Jacobi smoother, defined by

$$\begin{aligned} G_p(u_p, f_p) &= (I - 2\omega c_p A_p)u_p + 2\omega c_p f_p \\ &= \bar{G}_p u_p + (I - \bar{G}_p)A_p^{-1} f_p \end{aligned} \tag{2.5}$$

where $c_p = h_p^2/8$, $p = 1,\ldots, k$, and $\bar{G}_p$ is the linear part of G_p. We require that $0 < \omega < 1$. We do not allow $\omega = 1$, which would correspond to a Jacobi iteration. The constant, c_p, is approximately equal to the inverse of the spectral radius, $\rho(A_p)$. In fact, $c_p\rho(A_p) = 1 - 0(h_p^2)$, and therefore $\bar{G}_p$ is a contraction, i.e.,

$$\rho(I - 2\omega c_p A_p) < 1\,. \tag{2.6}$$

We define inner products and norms by:

$$\langle u^p, v^p \rangle_p = h_p^2 \sum_{x \epsilon \Omega_p} u^p(x) \bar{v}^p(x) \tag{2.7a}$$

and

$$\|u^p\|^2 = \langle u^p, u^p \rangle_p \ , \tag{2.7b}$$

for u^p, v^p defined on Ω^p .

For the projection and weighting operators we take I_{p-1}^p to be linear interpolation:

$$I_{p-1}^p = \frac{1}{4} \left] \begin{matrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{matrix} \right[_{h_{p-1}}^{h_p} \ , \tag{2.8a}$$

and I_p^{p-1} to be the adjoint of I_{p-1}^p relative to the discrete – L^2 inner products defined by (2.7a):

$$I_p^{p-1} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}_{h_p}^{h_{p-1}} \ , \tag{2.8b}$$

where we have used the "distribution" and "collection" stencils as in [10].

In the eigenfunction analysis we need some notation and simple formulas. Let $i = (i_1, i_2)$. Define

$$\xi_i^{(p)} = \cos^2 \left(\frac{i_1 \pi h_p}{2} \right) \tag{2.9a}$$

and

$$\eta_i^{(p)} = \cos^2 \left(\frac{i_2 \pi h_p}{2} \right) . \tag{2.9b}$$

A simple trigonometric identity gives us

$$\xi_i^{(p-1)} = (1 - 2\xi_i^{(p)})^2 \tag{2.10a}$$

and

$$\eta_i^{(p-1)} = (1 - 2\eta_i^{(p)})^2 . \tag{2.10b}$$

The eigenvalues of A_p can be written as

$$\nu_i^{(p)} = \frac{4(2 - \xi_i^{(p)} - \eta_i^{(p)})}{h_p^2}\,. \tag{2.11}$$

A simple calculation shows us that the effect of the projection on the eigenvectors of A_p can be expressed as[2]

$$I_p^{p-1}\alpha_i^{(p)} = \xi_i^{(p)}\eta_i^{(p)}\alpha_i^{(p-1)}\,. \tag{2.12}$$

The corresponding formulas for interpolation is

$$\begin{aligned} I_{p-1}^{p}\alpha_i^{(p-1)} &= \xi_i^{(p)}\eta_i^{(p)}\alpha_i^{(p)} - (1-\xi_i^{(p)})\eta_i^{(p)}\alpha_{(N_p-i_1,i_2)}^{(p)} \\ &\quad - \xi_i^{(p)}(1-\eta_i^{(p)})\alpha_{(i_1,N_p-i_2)}^{(p)} + (1-\xi_i^{(p)})(1-\eta_i^{(p)})\alpha_{(N_p-i_1,N_p-i_1)}^{(p)}\,. \end{aligned} \tag{2.13}$$

Note that eigenvectors of A_p are also eigenvectors of $\bar{G}_p$. The eigenvalue, $g_i^{(p)}$, of $\bar{G}_p$, corresponding to $\alpha_i^{(p)}$, is given by

$$g_i^{(p)} = 1 - 2\omega c_p \nu_i^{(p)}, \tag{2.14}$$

where the constants c_p are related by

$$c_{p-1} = 4c_p. \tag{2.15}$$

When we apply the multigrid algorithm, we transfer vectors to coarser grids. In the process we lose information. In this two-dimensional problem with an $(h\text{-}2h)$ grid structure the four (if $i_1 \neq N_p/2$ and $i_2 \neq N_p/2$) eigenvectors $\alpha_{(i_1,i_2)}^{(p)}$, $-\alpha_{(N_p-i_1,i_2)}^{(p)}$, $-\alpha_{(i_1,N_p-i_2)}^{(p)}$ and $\alpha_{(N_p-i_1,N_p-i_2)}^{(p)}$, defined on Ω^p, are indistinguishable on Ω^{p-1}. There are also $2N_p-3$ eigenvectors as defined on Ω^p which are indistinguishable from the null vector as defined on Ω^{p-1}. This phenomenon is what is referred to as aliasing.

This aliasing plays an important role in the analysis of the multigrid process and we introduce the following notation. Given two multi-indices $i = (i_1, i_1)$ and $j = (j_1, j_2)$, consider $\alpha_i^{(k)}$ and $\alpha_j^{(k)}$. If $\alpha_i^{(p)} = \pm\alpha_j^{(p)}$ then we write $i \sim j\ (p)$. If $\alpha_i^{(p)}$ and $\alpha_j^{(p)}$ are not linearly dependent then $i \not\sim j\ (p)$.

---

(2) In the cases where $|\, i \,| := \max(i_1, i_2) \geq 1/N_p$, one should replace $\alpha_i^{(p-1)}$ by its proper (unique) representation, $\alpha_{\tilde{i}}^{(p-1)}$, where $|\, \tilde{i} \,| < N_{p-1}$. However, Formula (2.12) is also correct in this form.

2.2 Intergrid Operator Identities

A multigrid cycle consists of smoothings and intergrid transfers. The smoother is applied to reduce the high frequency (rough) components of the error. The residual is transfered to a coarser grid where solving exactly for the error correction is less expensive. By solving and then interpolating this coarse grid correction back to the fine grid, the low frequency (smooth) components of the error are reduced. In the boundary value problem (2.1), the eigenfunctions are easily identifiable as rough or smooth, being products of sine functions. The same is true for the discrete operators, A_p, $1 \le p \le k$. To gain insight into the properties of the multigrid process we study the effect of a multigrid cycle on the eigenvectors of A_k.

Using formulas (2.12) and (2.13) it is clear that transferring $\alpha_i^{(p)}$ from Ω^p to Ω^{p-1} and then interpolating back, results in a linear combination of the four eigenvectors which alias from Ω^p to Ω^{p-1} . A 'smooth' eigenvector, i.e. $\xi_i^{(p)}$ and $\eta_i^{(p)}$ close to zero, picks up 'rougher' components. In the full k-grid problem where there are 4^{k-1} vectors aliasing from Ω^k to Ω^1 , keeping track of the aliasing is difficult. Fortunately, there are a few simplifying features. The second of the following three Lemmas, in particular, simplifies the analysis. Define

$$I_{p_2}^{p_1} = I_{p_1+1}^{p_1} I_{p_1+2}^{p_1+1} \cdots I_{p_2}^{p_2-1} \;, \quad 1 \le p_1 < p_2 \le k \;. \tag{2.16}$$

Lemma 2.1

If $j \sim i\ (n)$ and $j \not\sim i\ (n+1)$ for some $0 \le n < k$, then

$$\langle \alpha_i^{(p)} \,,\, I_k^p \alpha_j^{(k)} \rangle \;=\; \begin{cases} 0 & \text{if } n < p \le k; \\ \left(\prod_{m=p+1}^{n} \xi_i^{(m)} \eta_i^{(m)} \right) \langle \alpha_i^{(n)}, I_k^n \alpha_j^{(k)} \rangle_n \langle \alpha_i^{(p)}, \alpha_i^{(p)} \rangle_p & p \le n \end{cases} \tag{2.17}$$

Proof of Lemma 2.1

Let $j \sim i\ (n)$ and $j \not\sim i\ (n+1)$ for n, $0 \le n < k$.

For $p > n$, the orthogonality of the $\alpha_i^{(p)}$ gives

$$\langle \alpha_i^{(p)} \ , \ I_k^p \alpha_j^{(k)} \rangle_p = 0. \tag{2.18}$$

For $p \le n$ and $i \not\sim (0,0)$ (p),

$$\langle I_p^n \alpha_i^{(p)}, \alpha_i^{(n)} \rangle_n = \left(\prod_{m=p+1}^{n} \xi_i^{(m)} \eta_i^{(m)} \right) \langle \alpha_i^{(n)}, \alpha_i^{(n)} \rangle_n \tag{2.19}$$

$$= \left(\prod_{m=p+1}^{n} \xi_i^{(m)} \eta_i^{(m)} \right) \langle \alpha_i^{(p)}, \alpha_i^{(p)} \rangle_p \, .$$

Since $I_k^p = I_n^p \, I_k^n$, then

$$\langle \alpha_i^{(p)}, I_k^p \alpha_j^{(k)} \rangle_p = \langle I_p^n \alpha_i^{(p)}, I_k^n \alpha_j^{(k)} \rangle_n \, . \tag{2.20}$$

Using $j \sim i$ (n) and (2.19) gives

$$\langle \alpha_i^{(p)}, I_k^p \alpha_j^{(k)} \rangle_p = \left(\prod_{m=p+1}^{n} \xi_i^{(m)} \eta_i^{(m)} \right) \langle \alpha_i^{(n)}, I_k^n \alpha_j^{(k)} \rangle_n \langle \alpha_i^{(p)}, \alpha_i^{(p)} \rangle_p \, . \tag{2.21}$$

If $i \sim (0,0)$ (p) , then (2.21) is trivially true. ∎

Lemma 2.2

For any n , $1 \le n \le k$, and $i \not\sim (0,0)$ (n) ,

$$\sum_{j \sim i \ (n)} \mid \langle \alpha_i^{(n)}, I_k^n \alpha_j^{(k)} \rangle_n \mid = 1 \, . \tag{2.22}$$

Proof of Lemma 2.2

If $n = k$ then $j \sim i$ (1) implies $j = i$. Since $\langle \alpha_i^{(k)}, \alpha_i^{(k)} \rangle_k = 1$, (2.22) holds for $n = k$.

Assume $\sum_{j \sim i \ (s+1)} \mid \langle \alpha_i^{(s+1)}, I_k^{s+1} \alpha_j^{(k)} \rangle_{s+1} \mid = 1$ for s, where $s < k$.

Define

$$\begin{aligned} i^1 &= i = (i_1, i_2) \, , \\ i^2 &= (N_{s+1} - i_1, i_2) \, , \\ i^3 &= (N_{s+1} - i_1, N_{s+1} - i_2) \, , \\ i^4 &= (i_1, N_{s+1} - i_2) \, . \end{aligned} \tag{2.23}$$

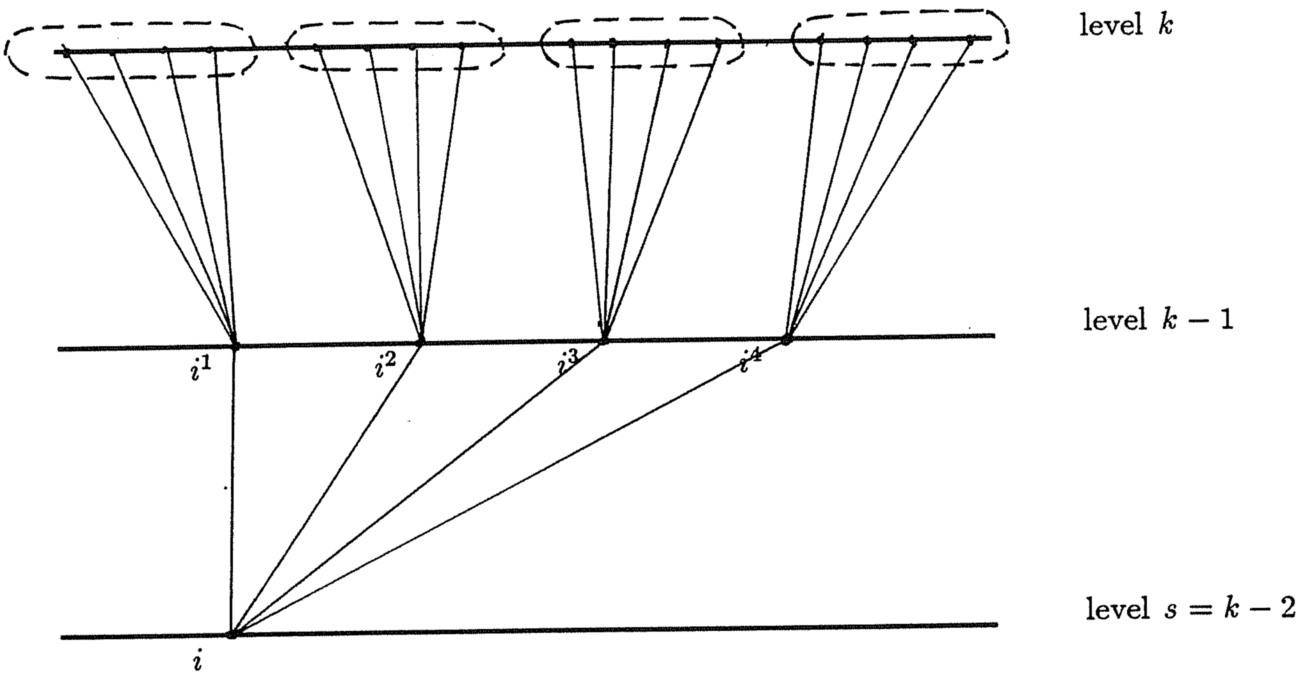

Figure 2.1: A splitting of the j, $j \sim i\ (s)$, where $s = k-2$.

The set $\{j \mid j \sim i\ (s)\}$ can be split into four disjoint subsets corresponding to all $j \sim i^1\ (s+1)$, $j \sim i^2\ (s+1)$, $j \sim i^3\ (s+1)$ and $j \sim i^4\ (s+1)$. Figure 2.1 shows this schematically for the case $s = k-2$. Therefore the summation can be split as:

$$\begin{aligned}\sum_{j \sim i\ (s)} \mid \langle \alpha_i^{(s)}, I_k^s \alpha_j^{(k)} \rangle_s \mid &= \sum_{j \sim i\ (s)} \mid \langle I_s^{s+1} \alpha_i^{(s)}, I_k^{s+1} \alpha_j^{(k)} \rangle_{s+1} \mid \\ &= \Bigg(\sum_{j \sim i^1\ (s+1)} + \sum_{j \sim i^2\ (s+1)} + \sum_{j \sim i^3\ (s+1)} + \sum_{j \sim i^4\ (s+1)} \Bigg) \\ &\qquad \mid \langle I_s^{s+1} \alpha_i^{(s)}, I_k^{s+1} \alpha_j^{(k)} \rangle_{s+1} \mid \ . \end{aligned} \tag{2.24}$$

Using (2.13) and the orthogonality of the $\alpha_i^{(s+1)}$, the summation can be written as:

$$\sum_{j\sim i\ (s)} |\langle\alpha_i^{(s)}, I_k^s\alpha_j^{(k)}\rangle_s| = \xi_i^{(s+1)}\eta_i^{(s+1)} \sum_{j\sim i^1\ (s+1)} |\langle\alpha_{i^1}^{(s+1)}, I_k^{s+1}\alpha_j^{(k)}\rangle_{s+1}|)$$
$$+ (1-\xi_i^{(s+1)})\eta_i^{(s+1)} \sum_{j\sim i^2\ (s+1)} |\langle\alpha_{i^2}^{(s+1)}, I_k^{s+1}\alpha_j^{(k)}\rangle_{s+1}|$$
$$+ (1-\xi_i^{(s+1)})(1-\eta_i^{(s+1)}) \sum_{j\sim i^3\ (s+1)} |\langle\alpha_{i^3}^{(s+1)}, I_k^{s+1}\alpha_j^{(k)}\rangle_{s+1}|$$
$$+ (\xi_i^{(s+1)})(1-\eta_i^{(s+1)}) \sum_{j\sim i^4\ (s+1)} |\langle\alpha_{i^4}^{(s+1)}, I_k^{s+1}\alpha_j^{(k)}\rangle_{s+1}| \ .$$

By the inductive hypothesis, each summation on the right hand side is equal to one and the coefficients also sum to one. ∎

Lemma 2.3

For all n, $1 \le n \le k$, and $i \not\sim (0,0)\ (n)$,

$$\sum_{\substack{j\sim i\ (n)\\ j\not\sim i\ (n+1)}} |\langle\alpha_i^{(n)}, I_k^n\alpha_j^{(k)}\rangle_n| = 1 - \xi_i^{(n+1)}\eta_i^{(n+1)} \tag{2.25}$$

Proof of Lemma 2.3

Identity (2.25) follows directly from Lemma 2.2 since

$$\sum_{\substack{j\sim i\ (n)\\ j\not\sim i\ (n+1)}} |\langle\alpha_i^{(n)}, I_k^n\alpha_j^{(k)}\rangle_n| = \sum_{j\sim i\ (n)} |\langle\alpha_i^{(n)}, I_k^n\alpha_j^{(n)}\rangle_n|$$
$$- \xi_i^{(n+1)}\eta_i^{(n+1)} \sum_{j\sim i\ (n+1)} |\langle\alpha_i^{(n+1)}, I_k^{n+1}\alpha_j^{(k)}\rangle_{n+1}|$$
$$= 1 - \xi_i^{(n+1)}\eta_i^{(n+1)} \ . \ ∎$$

3.1 Definition of the Preconditioner

The multigrid preconditioner is based on the discrete five point Laplacian. M_k is one standard multigrid symmetric V–cycle starting with zero as the initial guess, except that the coarse grid correction is obtained by smoothing instead of by solving exactly on the coarsest grid. Having choosen a fixed number of grids, k, the multigrid preconditioner is defined recursively. Choose a positive (integer) number of smoothings, r. Then $M_k f_k :=$ $\bar{u}_k$ where $\bar{u}_p$ $(= M_p f_p)$, for f_p defined on Ω^p, $p = 1, \ldots, k$, is given by:

1.) Smooth r times starting with initial guess $= 0$:

$$\tilde{u}_p = G_p^r(0, f_p) \,. \tag{3.1a}$$

2.) Compute the residual and transfer to the coarse grid:

$$r_p = f_p - A_p \tilde{u}_p, \; f_{p-1} = I_p^{p-1} r_p. \tag{3.1b}$$

3.) Compute the coarse grid correction:

$$\text{If} \quad p = 2, \; \bar{u}_{p-1} = \bar{u}_1 = G_1^{2r}(0, f_1) \tag{3.1c}$$

$$\text{If} \quad p > 2, \; \bar{u}_{p-1} = M_{p-1} f_{p-1}. \tag{3.1d}$$

4.) Add the coarse grid correction:

$$\hat{u}_p = \tilde{u}_p + I_{p-1}^p \bar{u}_{p-1}. \tag{3.1e}$$

5.) Smooth r times starting with initial guess $= \hat{u}_p$:

$$\bar{u}_p = G_p^r(\hat{u}_p, f_p) \,. \tag{3.1f}$$

Because we have started with an initial guess of zero, the multigrid preconditioner is a linear operator acting on f_k. This definition of M_k can be rewritten as:

$$M_p = \left(I - \bar{G}_p^{2r}\right) A_p^{-1} + \bar{G}_p^r I_{p-1}^p M_{p-1} I_p^{p-1} \bar{G}_p^r \quad p = 2, \ldots, k \tag{3.2}$$

and $$M_1 = \left(I - \bar{G}_1^{2r}\right) A_1^{-1} \,.$$

These identities rely on the commutivity of $\bar{G}_p$ and A_p, $p = 1, 2, \ldots, k$.

3.2 The Problem

As remarked in the introduction, it is sufficient to examine the effectiveness of the multigrid preconditioner by considering the model problem (1.3). We take $\Omega = (0,1) \times (0,1)$ and ε real and small. It is for this model operator, $A_L^\varepsilon = -\varepsilon^2 \triangle + I$, that we prove our basic results.

Define

$$A_h^\varepsilon = \varepsilon^2 A_k + I. \tag{3.3}$$

Writing the symmetric preconditioner as $M_k = Q_k^* Q_k$, the preconditioned system is $A_h^{\varepsilon\prime} v' = F'$ where $A_h^{\varepsilon\prime} = Q_k A_h^\varepsilon Q_k^*$. Experimental evidence suggests the following:

Conjecture:

Let $r > 0$, $0 < \omega < 1$, $h > 0$ and $\varepsilon > h$. Choose the number of grid levels, k, so that $h_1 = 2^{k-1} h \approx \varepsilon$. Define $M_h^\varepsilon = M_k$. Then there exist constants c_1, $c_2 > 0$ such that

$$c_1 \varepsilon^2 \;\le\; \lambda_{\min}(M_h^\varepsilon A_h^\varepsilon) \;\le\; \lambda_{\max}(M_h^\varepsilon A_h^\varepsilon) \;\le\; c_2 \varepsilon^2 .$$

What we prove in this paper is:

Theorem 3.1

Let $r = 1,2,3,4$ and $\omega = .7,.8,.9$ or $r = 2,3,4$ and $\omega = .5,.6$. Let $h \ge 1/8192$ and $\varepsilon > h$. Choose k so that $h_1 = 2^{k-1} h \approx \varepsilon$. Then there exist constants $c_1(h)$, $c_2(h) > 0$ such that

$$c_1(h)\varepsilon^2 \;\le\; \lambda_{\min}(M_h^\varepsilon A_h^\varepsilon) \;\le\; \lambda_{\max}(M_h^\varepsilon A_h^\varepsilon) \;\le\; c_2(h)\varepsilon^2 . \tag{3.4}$$

Remark 3.1

For fixed ε, r and ω, numerical evidence indicates that, as $h \to 0$,

$$c_1(h) \to c_1 > 0$$
$$c_2(h) \to c_2 > 0.$$

Remark 3.2:

Since $A_h^{\varepsilon\prime}$ is similar to $M_h^\varepsilon A_h^\varepsilon$, (3.4) implies that $K(A_h^{\varepsilon\prime})$ is bounded independent of ε.

Proof of Theorem 3.1:

Define

$$\mu_{ij} = \langle\, M_k\left(\varepsilon^2 A_k + I\right) \alpha_i^{(k)},\ \alpha_j^{(k)}\, \rangle_k. \tag{3.5}$$

Because of the aliasing, μ_{ij} can be nonzero for $j \neq i$. However if $i \not\sim j$ (1) (i.e. $\alpha_i^{(k)}$ and $\alpha_j^{(k)}$ are distinguishable on the coarsest grid) then $\mu_{ij} = 0$.

Choose $m = (m_1, m_2)$ where $|m| := \max(m_1, m_2) < N_k$.

Let $j_1,\ j_2, \ldots,\ j_{4^{k-1}}$ be some ordering of the $j \sim m(1)$.

We now define $\mathcal{M}_m$ to be a $4^{k-1} \times 4^{k-1}$ matrix given by

$$\left(\mathcal{M}_m\right)_{p,q} = \mu_{j_p j_q}. \tag{3.6}$$

We consider the subspaces

$$S_m := \text{linear span}\left(\left\{\alpha_j^{(k)}\ :\ j \sim m\ (1)\right\}\right), \tag{3.7}$$

where $|m| < N_k$. The S_m are orthogonal (with respect to the inner product defined by (2.7a)) subspaces and invariant under $M_k(\varepsilon^2 A_k + I)$. Therefore if we show that

$$c_1\varepsilon^2 \leq \lambda_{\min}\left(\mathcal{M}_m\right) \leq \lambda_{\max}\left(\mathcal{M}_m\right) \leq c_2\varepsilon^2 \tag{3.8}$$

for each m, then (3.4) will be proved.

By the Gershgorin theorem, any eigenvalue, λ, of $\mathcal{M}_m$ must satisfy

$$|\lambda - \mu_{ii}| \leq \sum_{\substack{j \sim i\ (1) \\ j \neq i}} |\mu_{ij}| \tag{3.9}$$

for some $i \sim m$ (1).

We show that $\mathcal{M}_m$ is diagonally row dominant and therefore we can use information about the behaviour of the diagonal entries of $\mathcal{M}_m$ to prove (3.8). Specifically, in Section 3.3 we give a computable formula, (3.22), for a quantity $C^i_{h,k,r,\omega}$, independent of ε, such that

$$\sum_{\substack{j \sim i\ (1) \\ j \neq i}} |\mu_{ij}| \leq C^i_{h,k,r,\omega}\, \mu_{ii}. \tag{3.10}$$

For certain choices of r and ω, $C^i_{h,k,r,\omega}$ has been computed, for every i, showing that $\bar{C}_{h,k,r,\omega} := \sup_i C^i_{h,k,r,\omega} < 1$ for the $k = 2, 3, \ldots, 12$ grid problems, using $h = 2^{-1}$ to $h = 2^{-13}$. See Section 3.4. In Section 3.5 it is shown that $\exists\ \underline{c}, \bar{c} > 0$ such that

$$\underline{c}\varepsilon^2 \leq \min_{|i| < N_k} \mu_{ii} \leq \max_{|i| < N_k} \mu_{ii} \leq \bar{c}\varepsilon^2. \tag{3.11}$$

Combining (3.9),(3.10) and (3.11) we have, for any eigenvalue, λ, of $\mathcal{M}_m$,

$$\left(1 - \bar{C}_{h,k,r,\omega}\right) \underline{c}\varepsilon^2 \leq \lambda \leq \left(1 + \bar{C}_{h,k,r,\omega}\right) \bar{c}\varepsilon^2, \tag{3.12}$$

which verifies (3.8) with $c_1 = \left(1 - \bar{C}_{h,k,r,\omega}\right) \underline{c}$ and $c_2 = \left(1 + \bar{C}_{h,k,r,\omega}\right) \bar{c}$.

Note that a common factor, $\varepsilon^2 v_i^{(k)} + 1$, appears in all the μ_{ij}, $j \sim i\ (1)$, therefore (3.10) is equivalent to

$$\sum_{\substack{j \sim i\ (1) \\ j \neq i}} \left| \langle M_k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k \right| \leq C^i_{h,k,r,\omega} \langle M_k \alpha_i^{(k)}, \alpha_i^{(k)} \rangle_k. \tag{3.13}$$

Let

$$D_i := \langle M_k \alpha_i^{(k)}, \alpha_i^{(k)} \rangle_k. \tag{3.14}$$

3.3 Bounds on the Off–Diagonal Elements of $\mathcal{M}_m$.

When applying a multigrid–type cycle to an eigenvector, $\alpha_i^{(k)}$, of A_k, the resulting vector, $M_k \alpha_i^{(k)}$, is a linear combination of $\alpha_i^{(k)}$ and all of the other eigenvectors, $\alpha_j^{(k)}$, which alias with $\alpha_i^{(k)}$ on the coarsest grid. In this section we give a formula for a bound on this aliasing. Specifically, we find an expression, $C^i_{h,k,r,\omega}$, where

$$J_i := \sum_{\substack{j \sim i\ (1) \\ j \neq i}} |\langle M_k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k| \leq C^i_{h,k,r,\omega} \langle M_k \alpha_i^{(k)}, \alpha_i^{(k)} \rangle_k. \tag{3.15}$$

Let $i = (i_1, i_2), h, k, r$ and ω be fixed.

Define

$$\xi_p = \cos^2\left(\frac{i_1 \pi h_p}{2}\right) \tag{3.16a}$$

$$\eta_p = \cos^2\left(\frac{i_2 \pi h_p}{2}\right) \tag{3.16b}$$

$$g_p = \langle\, \bar{G}_p^r \alpha_i^{(p)}, \alpha_i^{(p)} \,\rangle_p \tag{3.16c}$$

$$e_p = \langle \left(\sum_{\sigma=0}^{2r-1} \bar{G}_p^\sigma \right) \alpha_i^{(p)}, \alpha_i^{(p)} \,\rangle_p \tag{3.16d}$$

$$\text{and} \qquad \nu_p = \langle\, A_p \alpha_i^{(p)}, \alpha_i^{(p)} \,\rangle_p, \tag{3.16e}$$

where the i, r, h and ω dependence has been suppressed in the notation and only the grid level is displayed.

The following lemma gives a formula for any entry in the row of $\mathcal{M}_m$ corresponding to i, where $i \sim m\ (1)$.

Lemma 3.1

For any $j \sim i\ (1)$,

$$\langle M_k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k = 2\omega c_k \sum_{p=1}^{k} e_p \left(\prod_{m=p+1}^{k} 4 g_m \xi_m \eta_m \right) \langle \alpha_i^{(p)}, I_{p+1}^p G_{p+1}^r \cdots I_k^{k-1} G_k^r \alpha_j^{(k)} \rangle_p. \tag{3.17}$$

Proof of Lemma 3.1

A proof by induction shows that for every s, $2 \le s \le k$,

$$\langle M_s \alpha_i^{(s)}, \alpha_j^{(s)} \rangle_s = 2\omega c_s \sum_{p=1}^{s} e_p \left(\prod_{m=p+1}^{s} 4 g_m \xi_m \eta_m \right) \cdot \langle \alpha_i^{(s)}, I_{p+1}^p G_{p+1}^r \cdots I_s^{s-1} G_s^r \alpha_j^{(s)} \rangle_p. \tag{3.18}$$

Taking $s = k$ gives (3.17).

For $s = 2$, (3.4), (3.16) and (2.12) give

$$\begin{aligned} \langle M_2 \alpha_i^{(2)}, \alpha_j^{(2)} \rangle_2 &= \langle \left(I - G_2^{2r}\right) A_2^{-1} \alpha_i^{(2)}, \alpha_j^{(2)} \rangle_2 \\ &\quad + \langle \left(I - G_1^{2r}\right) A_1^{-1} I_2^1 G_2^r \alpha_i^{(2)},\ I_2^1 G_2^r \alpha_j^{(2)} \rangle_1 \qquad (3.19) \\ &= 2\omega c_2 e_2 \langle \alpha_i^{(2)}, \alpha_j^{(2)} \rangle_2 + 2\omega c_1 e_1 g_2 \xi_2 \eta_2 \langle \alpha_i^{(1)},\ I_2^1 G_2^r \alpha_j^{(2)} \rangle_1. \end{aligned}$$

Substituting $4c_2 = c_1$, proves (3.18) for $s = 2$.

Assume (3.18) is true for $s-1$ grids, $s>2$. For the s-grid problem, (3.4), (3.16) and (2.12) give

$$\begin{aligned}\langle M_s\alpha_i^{(s)},\alpha_j^{(s)}\rangle_s &= \langle (I-G_s^{2r})\,A_s^{-1}\alpha_i^{(s)},\alpha_j^{(s)}\rangle_s \\ &\quad + \langle M_{s-1}I_s^{s-1}G_s^r\alpha_i^{(s)},I_s^{s-1}G_s^r\alpha_j^{(s)}\rangle_{s-1} \\ &= 2\omega c_s e_s\langle\alpha_i^{(s)},\alpha_j^{(s)}\rangle_s \\ &\quad + \xi_s\eta_s g_s\langle M_{s-1}\alpha_i^{(s)},I_s^{s-1}G_s^r\alpha_j^{(s)}\rangle_{s-1}.\end{aligned} \tag{3.20}$$

Using the inductive hypothesis with $I_s^{s-1}G_s^r\alpha_j^{(s)}$ replacing $\alpha_j^{(s-1)}$, and using $4c_s=c_{s-1}$ proves (3.18). ∎

Lemma 3.1 can be used to get an expression for J_i, but the summation over all $j\sim i\ (1)$ would be difficult to compute. Theorem 3.2 shows that J_i can be bound by an expression which is no more complicated than the expression for $D_i=\langle M_k\alpha_i^{(k)},\alpha_i^{(k)}\rangle_k$.

We claim that the J_i can be bounded by an expression which is no more complicated than the expression for D_i $(=\langle M_k\alpha_i^{(k)},\alpha_i^{(k)}\rangle_k)$:

Theorem 3.2

a.) $$D_i = 2\omega c_k\sum_{p=1}^{k} e_p\left(\prod_{m=p+1}^{k} 4g_m^2\xi_m^2\eta_m^2\right). \tag{3.21a}$$

b.) $$J_i \le 2\omega c_k\sum_{p=1}^{k-1} e_p\left(1-\left(\prod_{m=p+1}^{k}\xi_m\eta_m\right)\right)\left(\prod_{m=p+1}^{k}4|g_m|\xi_m\eta_m\right). \tag{3.21b}$$

Proof of Theorem 3.2

a.) Using Lemma 3.1 with $j = i$, combined with equations (3.16c) and (2.12) proves (3.21a).

b.) To prove (3.21b), split the grid levels by partitioning the $j \sim i\ (1)$, $j \neq i$. See Figure 3.1 for a schematic illustration for $k = 3$. For each $n = 1, 2, \ldots, k-1$ consider the j's such that $j \sim i(n)$ but $j \not\sim i\ (n+1)$. Lemma 3.1, Lemma 2.1, Lemma 2.3 and (2.6) lead to the following bound:

$$J_i = \sum_{n=1}^{k-1} \sum_{\substack{j \sim i\ (n) \\ j \not\sim i\ (n+1)}} |\langle M_k^{-1} \alpha_i^{(k)},\, \alpha_j^{(k)} \rangle_k| \tag{3.22}$$

$$\leq 2\omega c_k \sum_{n=1}^{k-1} (1 - \xi_{n+1}\eta_{n+1}) \sum_{p=1}^{n} e_p \left(\prod_{m=p+1}^{n} \xi_m \eta_m \right) \left(\prod_{m=p+1}^{k} 4|g_m|\xi_m \eta_m \right).$$

Changing the order of summation gives

$$J_i \leq 2\omega c_k \sum_{p=1}^{k-1} e_p \left[\sum_{n=p}^{k-1} (1 - \xi_{n+1}\eta_{n+1}) \prod_{m=p+1}^{n} |g_m|\xi_m \eta_m \right] \tag{3.23}$$

$$\left(\prod_{m=p+1}^{k} 4|g_m|\xi_m \eta_m \right).$$

Observe that the quantity in square brackets can be simplified to:

$$1 - \prod_{m=p+1}^{k} \xi_m \eta_m. \quad \blacksquare$$

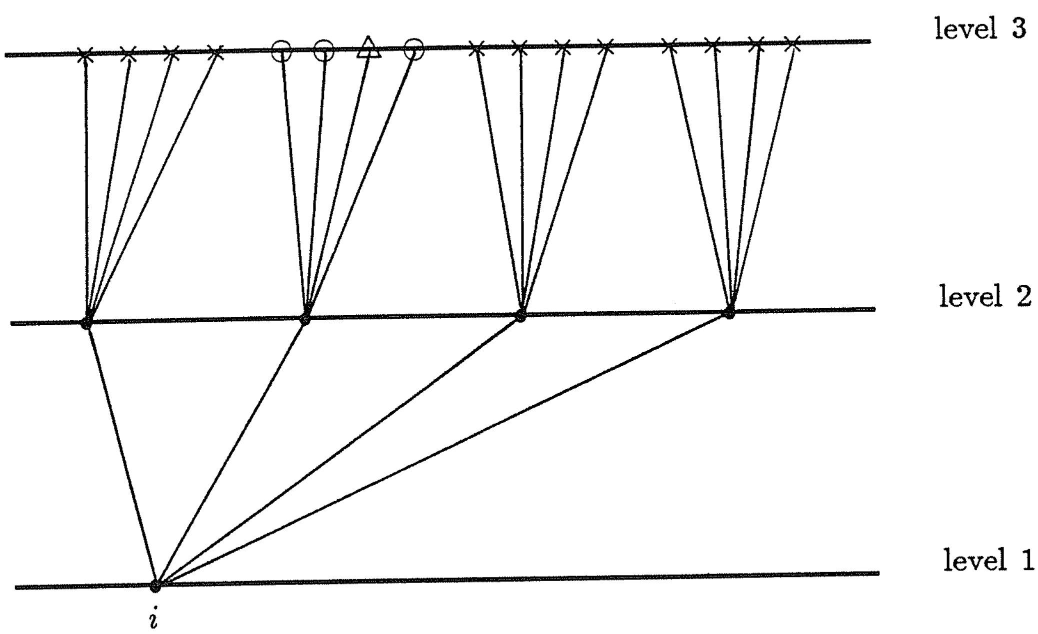

Figure 3.1: A splitting of the j, $j \sim i\ (s)$, $j \neq 1$.

$\times$	$j \sim i\ (1)\ ,\ j \not\sim i\ (2)$
$\odot$	$j \sim i\ (2)\ ,\ j \not\sim i\ (3)$
$\triangle$	$j \sim i\ (3)$.

Remark 3.3

The constants $\bar{C}_{h,k,r,\omega}$ can now be expressed as

$$\bar{C}_{h,k,r,\omega} = \sup_i \left(C^i_{h,k,r,\omega}\right),$$

where

$$C^i_{h,k,r,\omega} = \frac{\sum_{p=1}^{k-1} e_p \left(1 - \prod_{m=p+1}^{k} \xi_m \eta_m\right)\left(\prod_{m=p+1}^{k} 4|g_m|\xi_m\eta_m\right)}{\sum_{p=1}^{k} e_p \left(\prod_{m=p+1}^{k} 4g_m^2\xi_m^2\eta_m^2\right)}. \tag{3.24}$$

Note that the denominator has one more term in the sum than does the numerator.

3.4 Computed Values of the Off–Diagonal Bounds

Ideally, one would like to find analytic bounds for $C^i_{h,k,r,\omega}$, independent of i, h and k. On the other hand, bounds are easily computed for any given h, k, r and ω.

Figures 3.2–3.5 indicate the dependence of $C^i_{h,k,r,\omega}$ on $i = (i_1, i_2)$ for $h = 1/64$, $r = 1$, $\omega = .8$ and k=2,3,4 and 5 grids. The maximum is taken on the boundaries $i_1 = 1$ or $i_2 = 1$. Along the boundary $i_2 = 1$ there are 2^{k-2} relative maxima for the k–grid problem. (For all values of h, k, r and ω tried, the maximum of C^i was attained at $(1,\ i_2)$ and $(i_2,\ 1)$ for some i_2.) Figures 3.6–3.9 show the dependence on r for $k = 4$ grids.

Tables 3.1–3.8 give the calculated bounds, $\sup_{|i|<1/h}\left(C^i_{h,k,r,\omega}\right)$, for $\omega = .5$, .8 and $r = 1$,2,3 and 4. The multi–index at which the supremum was attained is listed below the bound.

To find bounds for $\omega = .8$ and $r = 1, 2, 3, 4$, independent of h and k, we used $h = 1/8192$ (which means $>$ 67 million points on the fine grid). These numbers are bounds for all $h > 1/8192$ and all k corresponding to these meshsizes. Observing the asymptotic behaviour leads one to believe that they are also bounds for all $h < 1/8192$ and any number of grids, k. See Tables 3.9–3.10.

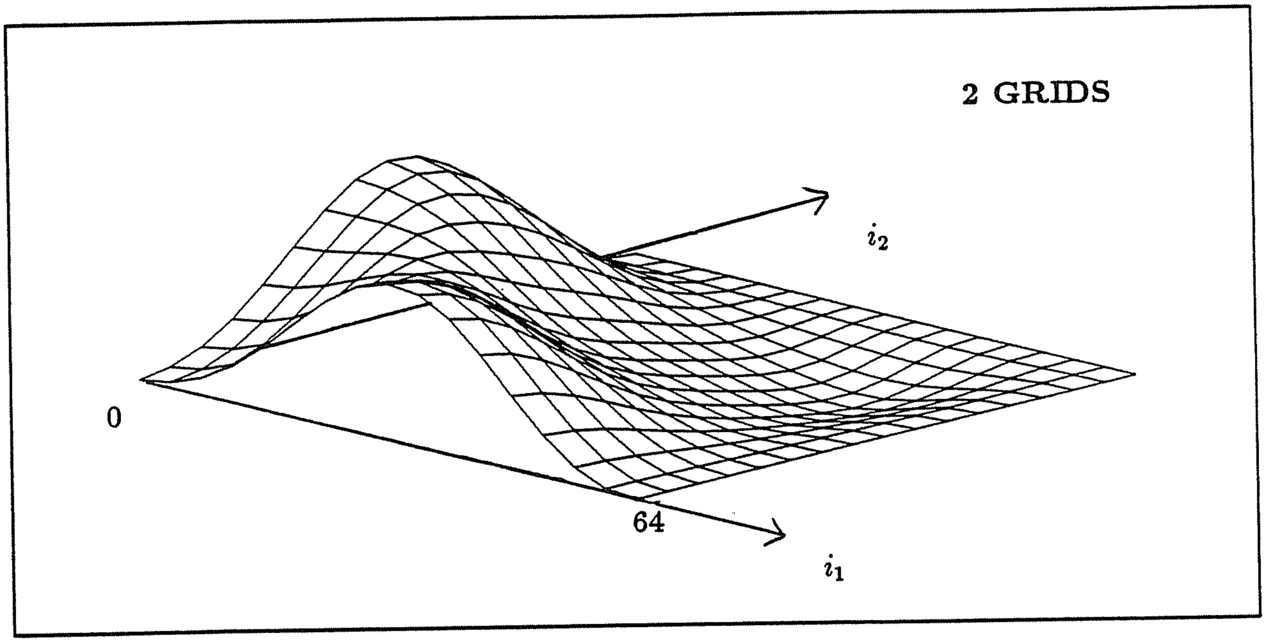

Figure 3.2: $C^i_{h,k,r,\omega}$ for $h = 1/64$, $r = 1$, $\omega = .8$

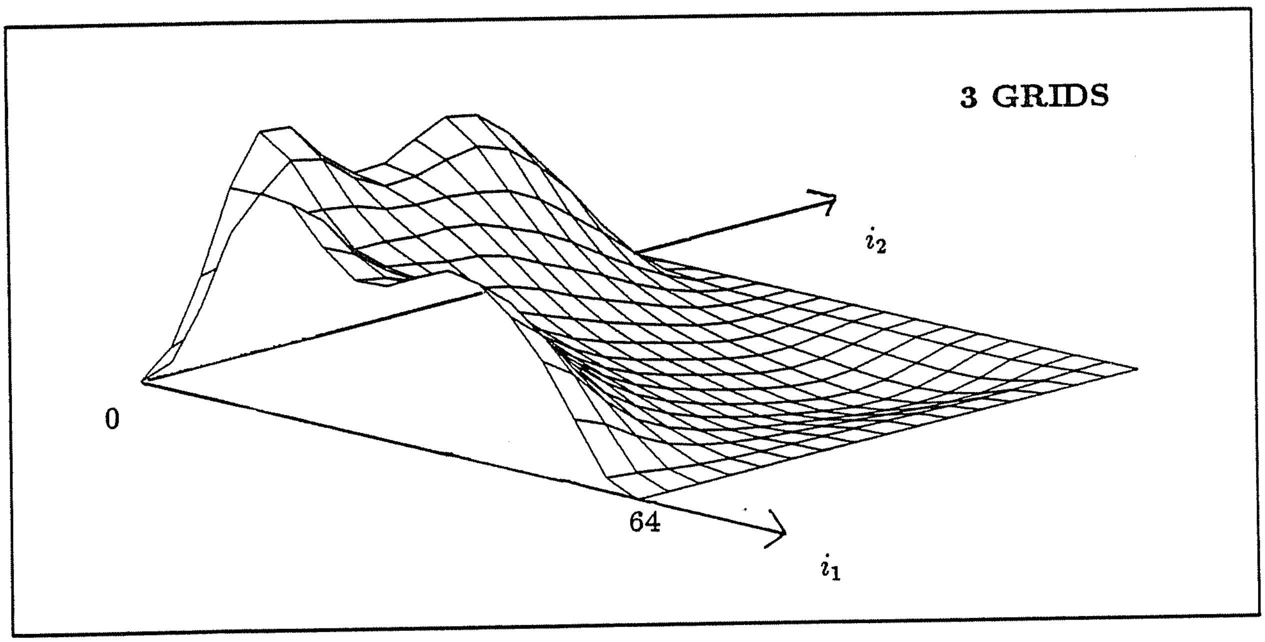

Figure 3.3: $C^i_{h,k,r,\omega}$ for $h = 1/64$, $r = 1$, $\omega = .8$

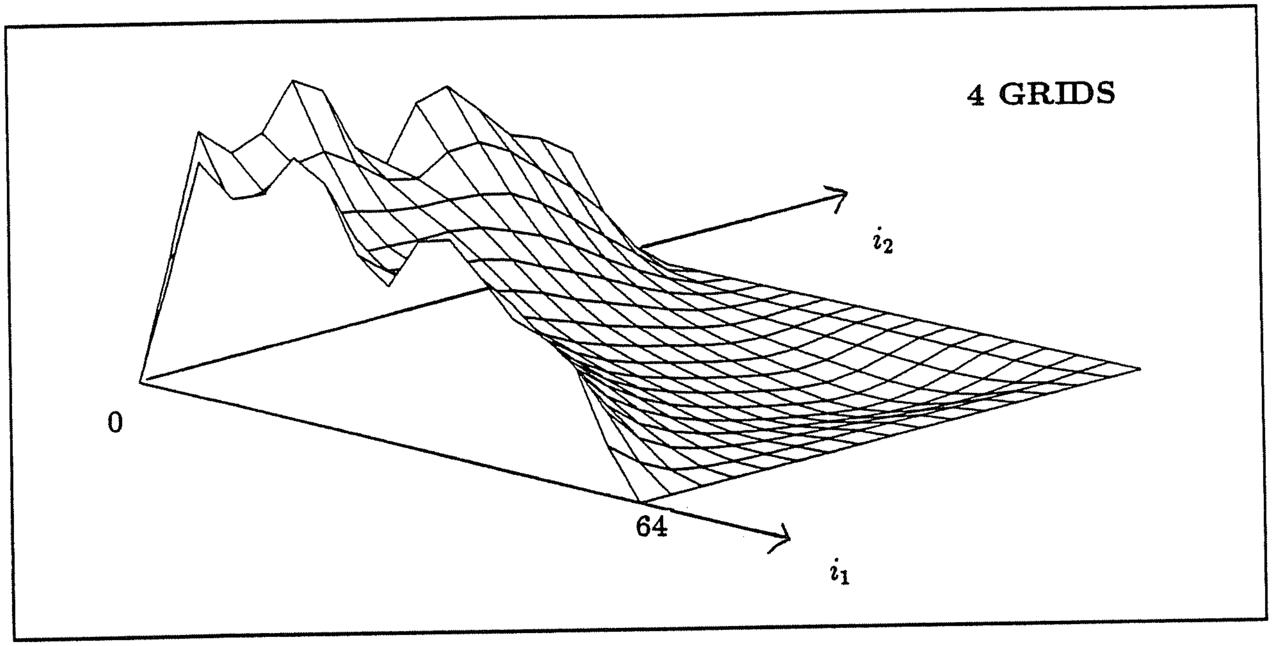

Figure 3.4: $C^i_{h,k,r,\omega}$ for $h = 1/64$, $r = 1$, $\omega = .8$

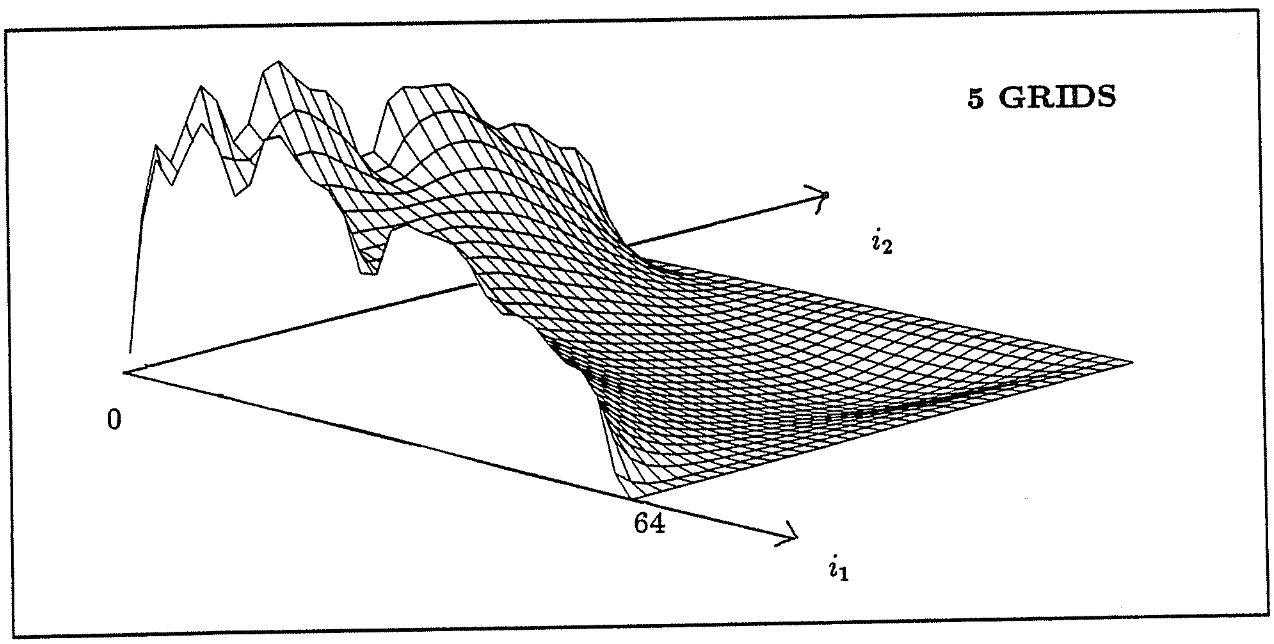

Figure 3.5: $C^i_{h,k,r,\omega}$ for $h = 1/64$, $r = 1$, $\omega = .8$

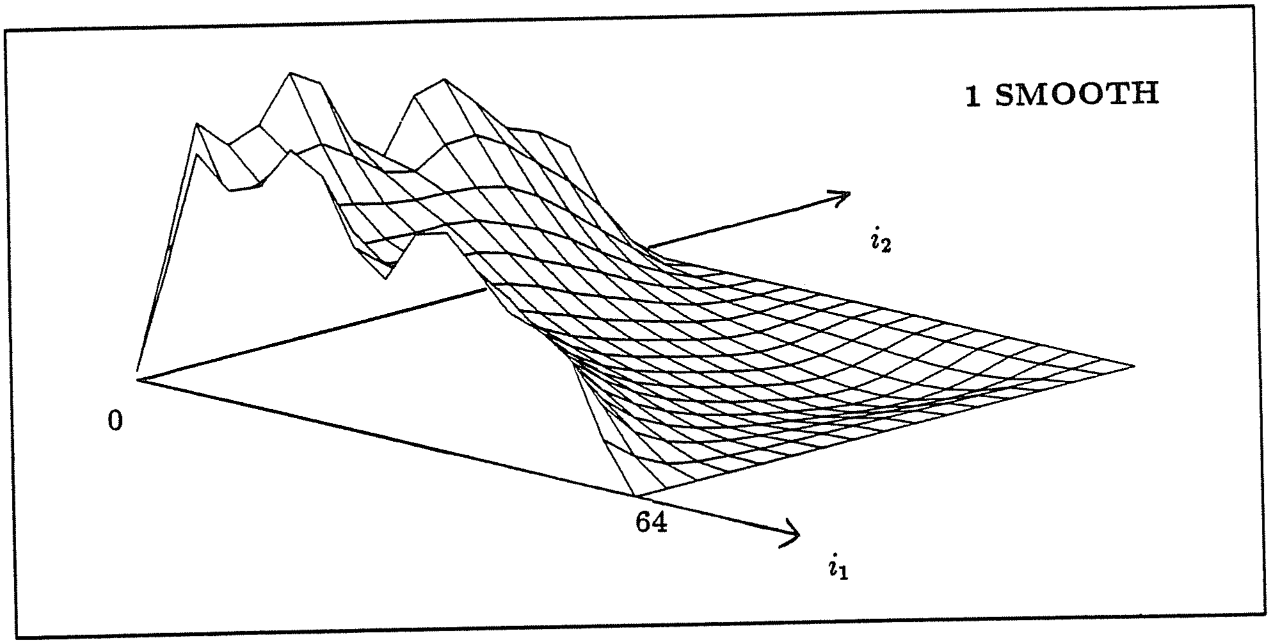

Figure 3.6: $C^i_{h,k,r,\omega}$ for $h = 1/64$, 4 grids, $\omega = .8$

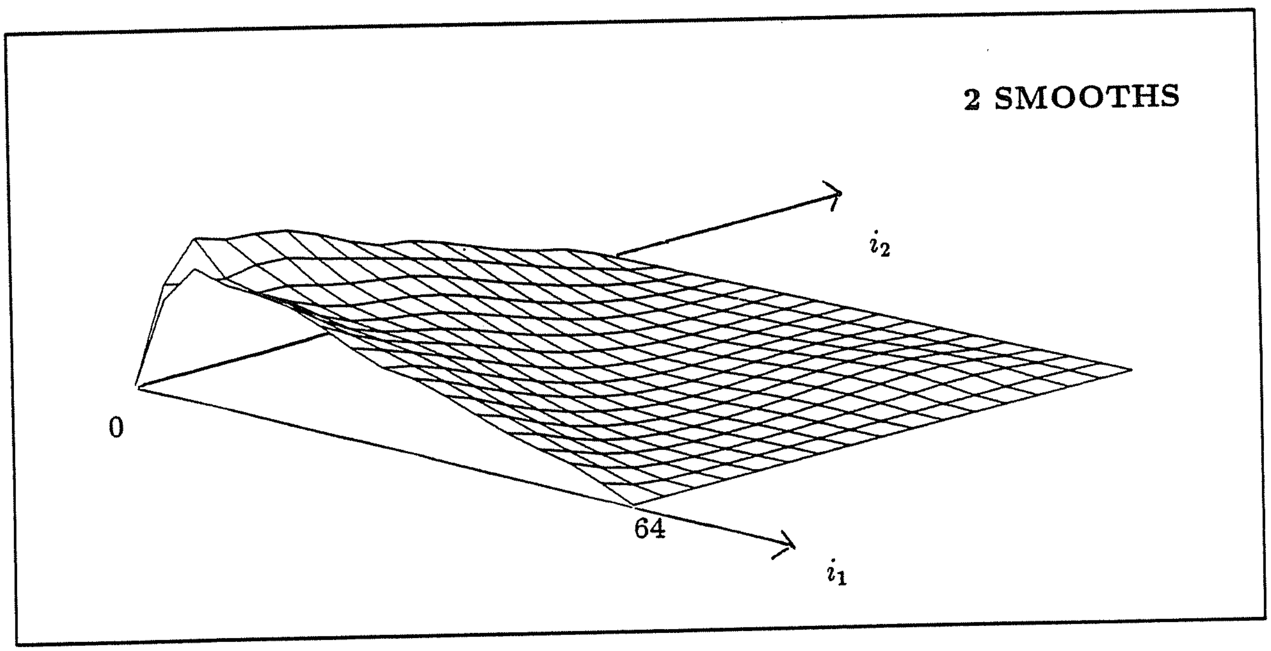

Figure 3.7: $C^i_{h,k,r,\omega}$ for $h = 1/64$, 4 grids, $\omega = .8$

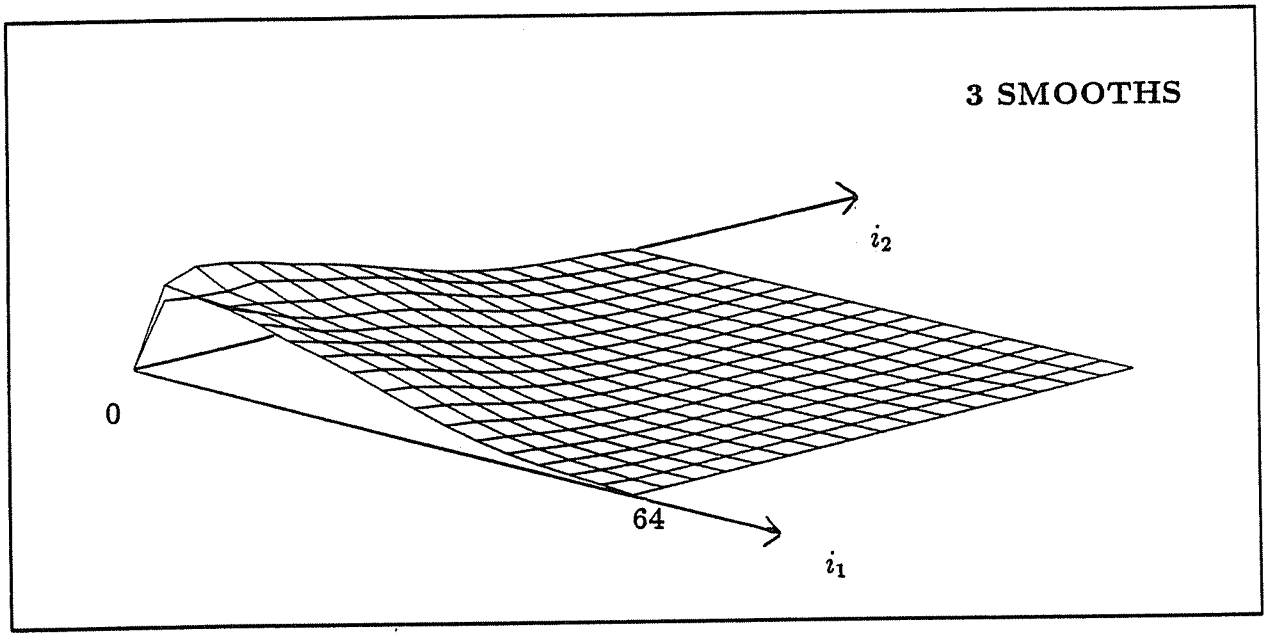

Figure 3.8: $C^i_{h,k,r,\omega}$ for $h = 1/64$, 4 grids, $\omega = .8$

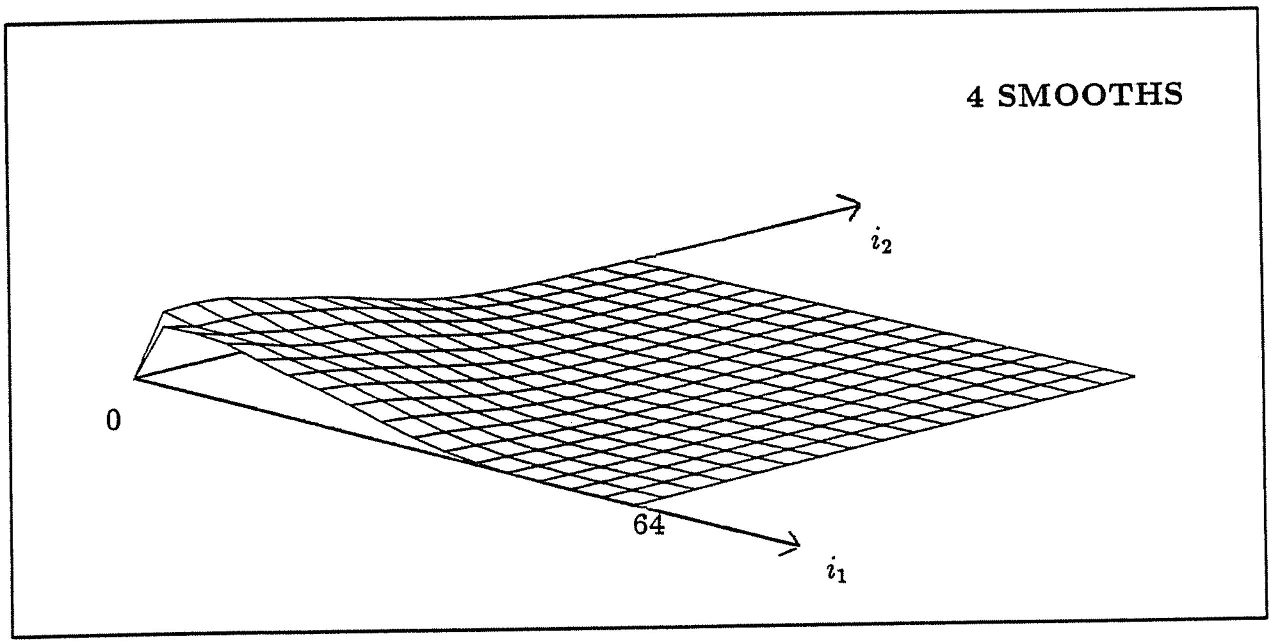

Figure 3.9: $C^i_{h,k,r,\omega}$ for $h = 1/64$, 4 grids, $\omega = .8$

Table 3.1 $\bar{C}_{h,k,r,\omega}$ $\omega = .5$, $r = 1$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.4640 (1,9)	.7165 (1,11)	.8026 (1,11)				
1/32	.4688 (1,19)	.7530 (1,22)	.9484 (1,21)	1.025 (1,11)			
1/64	.4707 (1,37)	.7632 (1,45)	.9942 (1,41)	1.149 (1,21)	> 1		
1/128	.4712 (1,74)	.7669 (1,89)	1.004 (1,81)	> 1	> 1	> 1	
1/256	.4712 (1,149)	.7669 (1,178)	1.006 (1,163)	> 1	> 1	> 1	> 1
1/512	.4712 (1,298)	.7671 (1,357)	1.007 (1,325)	> 1	> 1	> 1	> 1

Table 3.2 $\bar{C}_{h,k,r,\omega}$ $\omega = .5$, $r = 2$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	..3115 (1,9)	.4296 (1,5)	.4680 (1,5)				
1/32	.3196 (1,17)	.4574 (1,10)	.5441 (1,11)	.5573 (1,11)			
1/64	.3215 (1,35)	.4658 (1,19)	.5700 (1,22)	.6084 (1,21)	.6142 (1,11)		
1/128	.3220 (1,70)	.4680 (1,39)	.5771 (1,44)	.6277 (1,23)	.6591 (1,21)	.6643 (1,21)	
1/256	.3221 (1,139)	.4685 (1,77)	.5790 (1,88)	.6349 (1,45)	.6741 (1,41)	.6856 (1,43)	.6876 (1,43)
1/512	.3221 (1,278)	.4688 (1,155)	.5795 (1,177)	.6368 (1,91)	.6782 (1,82)	.6923 (1,86)	.6997 (1,43)

Table 3.3 $\bar{C}_{h,k,r,\omega}$ $\omega = .5$, $r = 3$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.2200 (1,7)	.2998 (1,5)	.3102 (1,5)				
1/32	.2271 (1,15)	.3283 (1,9)	.3484 (1,5)	.3581 (1,5)			
1/64	.2285 (1,31)	.3346 (1,18)	.3694 (1,10)	.3961 (1,11)	.3981 (1,11).		
1/128	.2289 (1,61)	.3362 (1,36)	.3756 (1,20)	.4093 (1,22)	.4165 (1,21)	.4170 (1,21)	
1/256	.2290 (1,123)	.3367 (1,71)	.3773 (1,39)	.4130 (1,44)	.4243 (1,23)	.4297 (1,21)	.4302 (1,21)
1/512	.2290 (1,246)	.3368 (1,142)	.3777 (1,78)	.4139 (1,88)	.4279 (1,46)	.4356 (1,41)	.4366 (1,42)

Table 3.4 $\bar{C}_{h,k,r,\omega}$ $\omega = .5$, $r = 4$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.1689 (1,7)	.2253 (1,3)	.2368 (1,3)				
1/32	.1732 (1,14)	.2481 (1,7)	.2659 (1,5)	.2684 (1,5)			
1/64	.1747 (1,27)	.2550 (1,15)	.2823 (1,9)	.2876 (1,10)	.2880 (1,9)		
1/128	.1749 (1,54)	.2567 (1,31)	.2868 (1,18)	.2945 (1,20)	.3013 (1,11)	.3016 (1,11)	
1/256	.1750 (1,109)	.2570 (1,63)	.2880 (1,35)	.2971 (1,19)	.3081 (1,23)	.3089 (1,19)	.3090 (1,19)
1/512	.1750 (1,217)	.2571 (1,126)	.2883 (1,70)	.2982 (1,38)	.3099 (1,47)	.3125 (1,25)	.3131 (1,26)

Table 3.5 $\bar{C}_{h,k,r,\omega}$ $\omega = .8$, $r = 1$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/8	.2801 (1,5)	.3473 (1,3)					
1/16	.3448 (1,9)	.4724 (1,5)	.5071 (1,5)				
1/32	.3565 (1,17)	.5080 (1,9)	.5894 (1,11)	.6015 (1,11)			
1/64	.3581 (1,34)	.5166 (1,19)	.6179 (1,22)	.6647 (1,21)	.6698 (1,21)		
1/128	.3586 (1,68)	.5188 (1,37)	.6256 (1,44)	.6857 (1,41)	.7243 (1,21)	.7293 (1,21)	
1/256	.3587 (1,136)	.5194 (1,75)	.6278 (1,87)	.6916 (1,45)	.7423 (1,41)	.7531 (1,43)	.7551 (1,43)
1/512	.3587 (1,273)	.5195 (1,150)	.6283 (1,175)	.6938 (1,89)	.7471 (1,82)	.7609 (1,86)	.7695 (1,43)

Table 3.6 $\bar{C}_{h,k,r,\omega}$ $\omega = .8$, $r = 2$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.1954 (1,7)	.2552 (1,3)	.2640 (1,3)				
1/32	.2001 (1,14)	.2840 (1,7)	.2993 (1,5)	.3013 (1,5)			
1/64	.2013 (1,28)	.2932 (1,15)	.3223 (1,9)	.3252 (1,10)	.3268 (1,9)		
1/128	.2016 (1,55)	.2956 (1,31)	.3285 (1,17)	.3337 (1,20)	.3387 (1,11)	.3389 (1,11)	
1/256	.2017 (1,111)	.2961 (1,63)	.3299 (1,34)	.3387 (1,18)	.3473 (1,21)	.3499 (1,19)	.3500 (1,19)
1/512	.2018 (1,222)	.2963 (1,127)	.3304 (1,69)	.3402 (1,36)	.3497 (1,42)	.3537 (1,39)	.3538 (1,38)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1/8192	.2018	.2963	.3305	.3407	.3505	.3550	.3581

Table 3.7 $\bar{C}_{h,k,r,\omega}$ $\omega = .8$, $r = 3$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.1353 (1,6)	.1880 (1,3)	.1891 (1,3)				
1/32	.1385 (1,12)	.1993 (1,7)	.2077 (1,3)	.2088 (1,3)			
1/64	.1393 (1,24)	.2032 (1,13)	.2233 (1,7)	.2240 (1,7)	.2243 (1,7)		
1/128	.1395 (1,47)	.2044 (1,27)	.2267 (1,15)	.2296 (1,7)	.2303 (1,7)	.2307 (1,7)	
1/256	.1396 (1,95)	.2046 (1,54)	.2278 (1,29)	.2339 (1,15)	.2341 (1,14)	.2353 (1,14)	.2353 (1,14)
1/512	.1396 (1,189)	.2047 (1,108)	.2281 (1,58)	.2348 (1,30)	.2357 (1,15)	.2369 (1,27)	.2370 (1,29)

Table 3.8 $\bar{C}_{h,k,r,\omega}$ $\omega = .8$, $r = 4$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.1039 (1,5)	.1441 (1,3)	.1482 (1,1)				
1/32	.1057 (1,10)	.1531 (1,6)	.1628 (1,3)	.1629 (1,3)			
1/64	.1067 (1,21)	.1557 (1,12)	.1705 (1,6)	.1716 (1,6)	.1716 (1,6)		
1/128	.1068 (1,42)	.1563 (1,24)	.1736 (1,13)	.1761 (1,7)	.1762 (1,6)	.1762 (1,7)	
1/256	.1069 (1,84)	.1565 (1,48)	.1742 (1,26)	.1788 (1,13)	.1795 (1,13)	.1797 (1,13)	.1797 (1,13)
1/512	.1069 (1,168)	.1565 (1,95)	.1744 (1,51)	.1795 (1,27)	.1803 (1,25)	.1808 (1,13)	.1810 (1,13)

Table 3.9 $\bar{C}_{h,k,r,\omega}$ $\omega = .8$, $h = 1/8192$

r	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids
1	.3587	.5196	.6284	.6945	.7487	.7638
2	.2018	.2963	.3305	.3407	.3505	.3550
3	.1396	.2047	.2282	.2351	.2370	.2375
4	.1069	.1566	.1745	.1798	.1812	.1818

r	8 grids	9 grids	10 grids	11 grids	12 grids	13 grids
1	.7800	.7896	.7933	.7953	.7948	.7951
2	.3581	.3589	.3590	.3592	.3592	.3952
3	.2390	.2394	.2398	.2398	.2398	.2398
4	.1821	.1824	.1825	.1825	.1825	.1825

Table 3.10 $\bar{C}_{h,k,r,\omega}$ $k = 12$, $h = 1/8192$

r	$\omega = .5$	$\omega = .6$	$\omega = .7$	$\omega = .8$	$\omega = .9$
1	> 1	> 1	.980	.795	.648
2	.721	.554	.439	.359	.305
3	.444	.345	.282	.240	.210
4	.318	.254	.212	.183	.161

3.5 Bounds on the Diagonal Elements of $\mathcal{M}_m$

Recall that the diagonal elements, μ_{ii}, of $\mathcal{M}_m$ where $i \sim m\ (1)$, are given by,

$$\mu_{ii} = \langle\, M_k A_h^{\varepsilon} \alpha_i^{(k)},\ \alpha_i^{(k)} \,\rangle_k. \tag{3.25}$$

Since $A_h^{\varepsilon} = \varepsilon^2 A_k + I$ and hence

$$\mu_{ii} = \left(\varepsilon^2 \nu_i^{(k)} + 1\right) D_i, \tag{3.26}$$

the bounds on the μ_{ii} can be obtained from suitable information about the D_i's. The following characterization of the effect of the preconditioner on smooth and rough eigenvectors of A_k is central to the analysis and was given by Goldstein in [7].

Theorem 3.3

For $r \geq 1$, ω suitably chosen and h sufficiently small, the D_i's are positive real numbers such that:

$$\text{a.)}\quad D_i = 0\left(h_1^2\right) \quad \text{for} \quad \nu_i^{(k)} < d/h_1^2 \tag{3.27a}$$

$$\text{b.)}\quad D_i = \frac{(1-\eta)}{\nu_i^{(k)}} \quad \text{for} \quad \nu_i^{(k)} \geq d/h_1^2 \tag{3.27b}$$

where $0 < \eta < 1$ and η is independent of h and d is a constant.

We prove a more explicit version of the same result:

Theorem 3.4

For $r \geq 1$, $0 < \omega < 1$ and a fixed constant, d, where $\frac{1}{2} < d \leq 2$,

$$\text{a.)}\quad \frac{\omega(1-\omega)}{\max\left(2, d(1+r\omega)\right)} h_1^2 \leq D_i \leq \frac{2r\omega}{3} h_1^2 \quad \text{for} \quad \nu_i^{(k)} < \frac{d}{h_1^2} \tag{3.28a}$$

$$\text{b.)}\quad \frac{d\omega(1-\omega)}{8(1+r\omega)\nu_i^{(k)}} \leq D_i \leq \frac{1}{\nu_i^{(k)}} \quad \text{for} \quad \nu_i^{(k)} \geq \frac{d}{h_1^2}. \tag{3.28b}$$

Proof: in appendix.

These theorems give us bounds on the μ_{ii}, and, for example, Theorem 3.4 leads to the following bounds:

$$\text{For} \quad \nu_i^{(k)} < \frac{d}{h_1^2}, \qquad \frac{\omega(1-\omega)h_1^2}{\max(2, d(1+r\omega))} \le \mu_{ii} \le \frac{2r\omega d}{3}\left(\varepsilon^2 + \frac{h_1^2}{d}\right) \tag{3.29a}$$

$$\text{For} \quad \nu_i^{(k)} \ge \frac{d}{h_1^2}, \qquad \frac{d\omega(1-\omega)\varepsilon^2}{8(1+r\omega)} \le \mu_{ii} \le \varepsilon^2 + \frac{h_1^2}{d}. \tag{3.29b}$$

Therefore, taking $h_1 \cong \varepsilon$, we prove (3.12).

Using the diagonal dominance of the matrices, $\mathcal{M}_m$, we can estimate the dependence of the condition number of $M_k(\varepsilon^2 A_k + I)$ on the ratio $\alpha = h_1^2/\varepsilon^2$ from the behaviour of the diagonal elements, μ_{ii}. From the inequalities (3.29) we get an estimate for the choice of α which minimizes the condition number:

$$\alpha_{\text{optimal}} = \frac{1}{8}\left(\frac{3}{2r\omega}\right)^2. \tag{3.30}$$

This predicts that the optimal number of grids decreases as the quantity $r\omega$ increases. One can also use (3.29) to show that it is better to choose too many grids, ($\alpha > \alpha_{opt}$), rather than too few, ($\alpha < \alpha_{opt}$), (see Figure 3.10). These observations all accurately describe the experimental results — see the next section.

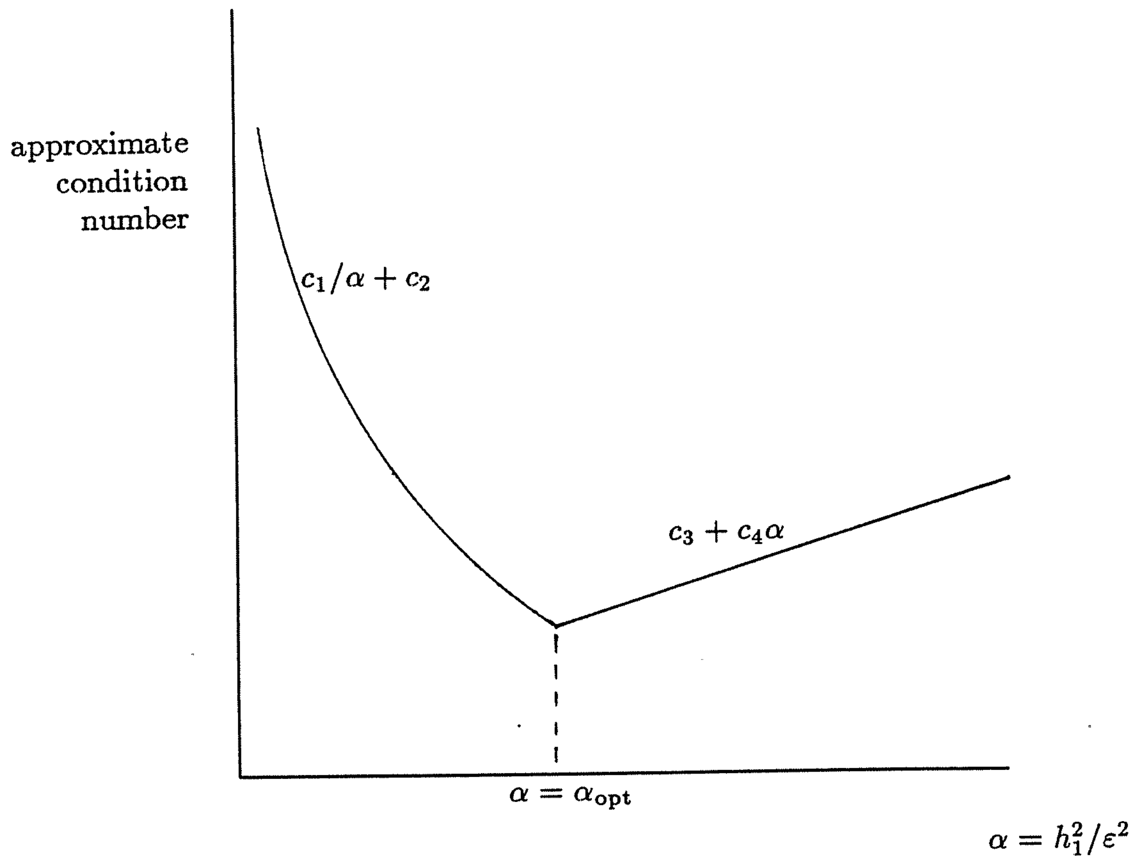

Figure 3.10: The condition number estimated from the diagonal terms.

4. Multigrid Preconditioner — Experimental Results.

Our numerical computations were carried out with three objectives in mind:

i) Observe the optimality of taking the meshsize on the coarsest grid, h_1 , to approximate the singular perturbation parameter, ε.

ii) Check the boundedness of the condition number of the multigrid–preconditioned system as ε and the fine grid meshsize, h, decrease.

iii) Compare the efficiency to other fast solvers, in particular, the corresponding multigrid algorithm used as an iterative solver.

We discretize the boundary value problem:

$$\begin{cases} A_L^\varepsilon u := (-\varepsilon^2 \triangle + I)u = f \quad \text{in } \Omega = (0,1)\times(0,1) \\ u = 0 \ \text{ on } \partial\Omega, \end{cases} \tag{4.1}$$

on a grid of uniform meshsize, h, as in Section 2.1. Using the multigrid preconditioner, M_h^ε, as defined in Section 3.1, we iteratively solve the discrete problem using a preconditioned conjugate gradient algorithm. Recall that k is the number of grids used in the multigrid algorithm, $h_k = h$, and the smoothers, G_p , $1 \le p \le k$,used to define M_h^ε, depend on the damping parameter, ω, and a fixed number of smooths per iteration, r. We solve

$$(\varepsilon^2 A_k + I)u_k = F_k, \tag{4.2}$$

starting with initial guess, u_k^0. We call this iterative solver PCCG(-$\triangle$,sm). The "$\triangle$" reminds us that the multigrid preconditioner is based on A_k, the negative of the discrete Laplacian, and not on the operator $A_k^\varepsilon = \varepsilon^2 A_k + I$ and "sm" indicates that we smooth instead of solving exactly on the coarsest grid. Experimentally, we find that a reasonably good choice of r and ω is $r = 2$ and $\omega = .8$ ($\omega = .8$ is optimal for the corresponding 2-grid multigrid solver, see [12]).

We first consider solving (4.2) with $F_k \equiv 1$. For $h = 1/64$ we show the dependence of the number of iterations required to reduce the norm of the residual by a factor of 10^{-6} on the choice of ε and h_1. See Table 4.1 . For given ε and h, the number of iterations listed is the largest observed for various choices of u_k^0. Note, in particular, the cases where $h_1 = \varepsilon$.

Table 4.2 displays the number of iterations required to reduce the relative error by a factor of 10^{-6} for various choices of h and ε, taking $h_1 = \varepsilon$. Here we used $F_k \equiv 0$.

Finally, we compare the efficiency of PCCG($-\Delta$,sm) to other elliptic solvers. We take $h = 1/64$, $\varepsilon = 1/8$, $F_k \equiv 1$ and an initial guess consisting of a smooth and a rough component, namely:

$$u_k^0 = 10 + 20\cos(64\pi x)\cos(64\pi y).$$

We consider a symmetric V-cycle, which is a fast iterative solver for equation (4.1), where we solve exactly on the coarsest grid (we use a symmetric band solver to invert $\varepsilon^2 A_1 + I$). We denote this algorithm by MULT. For comparison, an (extreme) choice of a preconditioner for the preconditioned conjugate gradient algorithm is considered, where the preconditioner is based on A_k^ε instead of A_k and we solve exactly on the coarsest grid. In other words, this preconditioner consists of *one cycle* of the solver, MULT, starting with *initial guess of zero.* This algorithm is called PCCG($-\varepsilon^2\Delta + I$,so). Of course we expect the behaviour of this preconditioner to be better than that of the simpler ($-\Delta$,sm) preconditioner, but we have the added expense of a coarse grid solve and (slightly) more complicated operator. Of interest to us here is that PCCG($-\varepsilon^2\Delta + I$,so) is not a significant improvement over PCCG($-\Delta$,sm) if the optimal choice of the number of grids is used.

In a conjugate gradient algorithm, the error reduction factor, $\|e_k\|/\|e_{k-1}\|$, typically decreases as k increases, whereas for a multigrid algorithm the error reduction factor increases as k increases. Therefore the preconditioned conjugate gradient routines will be more competitive when a large reduction in the relative residual is required and the multigrid algorithm is more competitive when a smaller reduction in the relative residual is required.

We also observe that increasing the number of smoothings per grid level will improve the performance of MULT more than it will improve the performance of the PCCG($-\Delta$,sm) algorithm. Similarly, optimizing the choice of the damping parameter, ω, will improve MULT more than it will improve PCCG($-\Delta$,sm).

Furthermore, one should keep in mind that, though it is difficult to improve the behaviour of the multigrid preconditioner, it is quite obvious how to improve the multigrid solver. Using better smoothers, or using a full multigrid algorithm (FMG) will dramatically improve the convergence rate.

Our first comparison is made with parameters which should give the PCCG($-\Delta$,sm) algorithm an advantage. We therefore consider a relatively inefficient choice of the damping parameter, $\omega = .5$, and require the norm of the residual to be reduced by a factor of 10^{-12}. The total cpu time (seconds) is recorded in Figure 4.3, with the number of iterations given in parentheses next to the time. The PCCG(-Δ,sm) algorithm appears to be competitive with MULT, at least for this meshsize, h. The PCCG(-$\varepsilon^2\Delta + I$,so) algorithm is only slightly faster.

We then take a more reasonable value of $\omega = .8$ and require the norm of the residual to be reduced by a factor of 10^{-6}. The total cpu time is recorded in Figure 4.4. The multigrid solver, MULT, is now the best choice.

All computations were done on a VAX 11/780.

We end this section with a few comments on the choice of using multigrid by itself as a solver, or using multigrid (based on a simpler operator) as a preconditioner:

- For the model problem (8.1), our experiments indicate that, for modest values of h and ε, a good multigrid algorithm is more efficient than a multigrid–preconditioned conjugate gradient algorithm.

- In a true variable coefficient problem, (1.1), the multigrid preconditioner has the advantage of being based on a constant coefficent operator. In this case, using multigrid as a preconditioner should be more competitive than in the model problem case. It is doubtful whether the multigrid preconditioner could outperform a good multigrid solver even in this case, but more testing would need to be done.

- In an indefinite problem, where multigrid solvers are more troublesome, one of the preconditioned conjugate gradient routines for indefinite problems might be preferable.

Table 4.1 Optimality of choosing $h_1 \approx \varepsilon$.
Largest (observed) # of iterations required for $\|r_k^i\|/\|r_k^0\| < 10^{-6}$.

$$F_k \equiv 1 \ , \ \omega = .8 \ , \ r = 2$$

h_1	$\varepsilon = 1/2$	$\varepsilon = 1/4$	$\varepsilon = 1/8$
1/32	> 20	> 20	20
1/16	12	12	10
1/8	9	8	8
1/4	7	7	9
1/2	7	8	9

Table 4.2 Boundedness of condition number independent of h and ε taking $\varepsilon = h_1$.
Largest (observed) # of iterations required for $\|u_k - u_k^i\|/\|u_k - u_k^0\| < 10^{-6}$.

$$F_k \equiv 0 \ , \ \omega = .8 \ , \ r = 2$$

h	$\varepsilon = 1/4$	$\varepsilon = 1/8$	$\varepsilon = 1/16$	$\varepsilon = 1/32$
1/32	5	6		
1/64	6	6	6	
1/128	6	6	6	6

Table 4.3 Experimental comparisons of approximate cpu time (sec).
Approximate cpu time (no. of iterations) required for $\|res_k\|/\|res_0\| < 10^{-12}$.

$$F_k \equiv 1 \ , \ \omega = .5 \ , \ r = 2$$

$$\varepsilon = 1/8 \quad , h = 1/64 \quad , u_k^0 = 10 + 20 \cdot \cos 64\pi x \cos 64\pi y$$

# of grids	MULT:V(2,2)	PCCG(-Δ,sm)	PCCG($-\varepsilon^2\Delta + I$,so)
2	61.3 (20)	– (>20)	53.4 (10)
4	44.2 (21)	40.6 (11)	39.2 (10)
6	44.4 (21)	44.8 (12)	39.5 (10)

Table 4.4 Experimental comparisons of approximate cpu time (sec).
Approximate cpu time (no. of iterations) required for $\|res_k\|/\|res_0\| < 10^{-6}$.

$$F_k \equiv 1 \ , \ \omega = .8 \ , \ r = 2$$

$$\varepsilon = 1/8 \quad , h = 1/64 \quad , u_k^0 = 10 + 20 \cdot \cos 64\pi x \cos 64\pi y$$

# of grids	MULT:V(2,2)	PCCG(-Δ,sm)	PCCG($-\varepsilon^2\Delta + I$,so)
2	24.3 (6)	49.9 (14)	35.2 (5)
4	14.3 (6)	22.4 (5)	29.6 (5)
6	14.4 (6)	23.8 (6)	29.7 (5)

5.1 V-cycle Convergence Bounds

In this section we briefly describe the results of applying the same techniques, in particular Lemma 2.2, to obtain bounds on the asymptotic convergence rates for multigrid V-cycles used to solve the Dirichlet problem for Poisson's equation in the unit square. The analysis is simpler in this case because we don't need diagonal dominance. Instead, we numerically evaluate the $\|\cdot\|_{\ell_\infty}$ norm of the appropriate matrix (i.e., the largest row sum of absolute values) which is a bound on the spectral radius. We present the details of this analysis in Section 5.2. We first define our basic multigrid V-cycle applied to the linear system

$$B_k U_k = F_k \tag{5.1}$$

starting with initial guess, u_k^0, with auxiliary problems, $B_p U_p = f_p$, $p = 1, 2, \ldots, k-1$, corresponding to discretizations on the coarser grids.

1. Initialize:

$$f_k \leftarrow F_k$$

$$u_k \leftarrow u_k^0$$

2. Update:

$$u_k \leftarrow \bar{u}_k$$

where each $\bar{u}_p$, $p = 2, 3, \ldots, k$ is defined recursively by:

(a.) Smooth r times starting with initial guess $= u_p$:

$$\tilde{u}_p = G_p^r(u_p, f_p) \tag{5.2a}$$

(b.) Compute the residual and transfer to the next coarser grid:

$$r_p = f_p - B_p \tilde{u}_p, \quad f_{p-1} = I_p^{p-1} r_p \tag{5.2b}$$

(c.) If $p > 2$ then return to (a.) to evaluate $\bar{u}_{p-1}$. If $p = 2$ then:

$$\bar{u}_1 = B_1^{-1} f_1 \tag{5.2c}$$

(d.) Add the coarse grid correction:

$$\hat{u}_p = \tilde{u}_p + I_{p-1}^p \bar{u}_{p-1} \tag{5.2d}$$

(e.) Smooth s times starting with initial guess $= \hat{u}_p$:

$$\bar{u}_p = G_p^s(\hat{u}_p, f_p) \tag{5.2e}$$

For the model problem analysis, we take Ω_p, A_p, I_{p-1}^p, I_p^{p-1} and $\bar{G}_p$ as defined in Section 2.1.

5.2 Error Analysis

Bounds on the asymptotic convergence factors of the multigrid cycles $M\backslash_{h,k,r,\omega}$ can be found in the following manner. Let $\varepsilon_k = U_k - u_k$ be the initial error and $\bar{\varepsilon}_k = U_k - \bar{u}_k$ be the error after one multigrid cycle, where U_k satisfies $A_k U_k = f_k$. In terms of the errors, definition (5.2) becomes:

(a) For $p = k, k-1, \cdots, 2$

$$\tilde{\varepsilon}_p = G_p^r \varepsilon_p ,$$

$$\varepsilon_{p-1} = A_{p-1}^{-1} I_p^{p-1} A_p \tilde{\varepsilon}_p .$$

(b) For p=1

$$\bar{\varepsilon}_1 = 0 .$$

(c) For $p = 2, \cdots, k$

$$\bar{\varepsilon}_p = \tilde{\varepsilon}_p - I_{p-1}^p(\varepsilon_{p-1} - \bar{\varepsilon}_{p-1}) .$$

Recall that $\bar{G}_p$ is the linear part of G_p. If $M^k \varepsilon_k = \bar{\varepsilon}_k$, then M^k is defined recursively by:

$$M^p = \bar{G}_p^r - I_{p-1}^p(I - M^{p-1}) A_{p-1}^{-1} I_p^{p-1} \bar{G}_p^r A_p , \quad 2 < p \le k \tag{5.3a}$$

$$M^1 = 0 . \tag{5.3b}$$

Note that the $\alpha_i^{(k)}$ are eigenvectors of A_k and G_k , but not of M^k . Define

$$S_i = \text{ linearspan } \{\alpha_j^{(k)} : j \sim i \ (1)\} . \tag{5.4}$$

By formulas (2.12) and (2.13) we see that the S_i are orthogonal subspaces which are invariant under M^k . Therefore a basis of eigenvectors, $\{v_\mu\}$, of M^k exists such that each v_μ can be written as

$$v_\mu = \sum_{j \sim i \ (1)} a_{j\mu} \alpha_j^{(k)} , \tag{5.5}$$

for some $i, | i |< N_k$, where $a_{j\mu} \in \mathbb{R}$. Since the $\alpha_i^{(k)}$ are orthonormal with respect to the discrete L_2 inner product, then

$$
\begin{aligned}
\langle M^k v_\mu, v_\mu \rangle_k &= \langle \sum_{j\sim i\ (1)} a_{j\mu} M^k \alpha_j^{(k)} \, , \sum_{m\sim i\ (1)} a_{m\mu} \alpha_m^{(k)} \rangle_k \\
&= \sum_{m\sim i\ (1)} a_{m\mu} \sum_{j\sim i\ (1)} a_{j\mu} \langle M^k \alpha_j^{(i)}, \alpha_j^{(k)} \rangle_k = \lambda_\mu \sum_{n\sim i\ (1)} a_{n\mu}^2
\end{aligned} \tag{5.6}
$$

where λ_μ is the eigenvalue of M^k corresponding to v_μ .

A bound on the λ_μ's will be a bound on the asymptotic convergence rate of the multigrid cycle. Let $\mathcal{M}_i$ be the $4^{k-1} \times 4^{k-1}$ matrix with $(\mathcal{M}_i)_{p,q} = \langle M^k \alpha_{j_p}^{(k)}, \alpha_{j_q}^{(k)} \rangle_k$ with $j_1, j_2, \cdots j_{4^{k-1}}$ some ordering of all the $j \sim i\ (1)$.

Remark 5.1 Note that for some i's, these j_p's are not necessarily unique. For example, if $i = (N_k/2, 1)$ then $(N_k/2, 1) = (N_k - N_k/2, 1)$.

Remark 5.2 The diagonal elements of $\mathcal{M}_i$ are the Rayleigh quotients,

$$
\frac{\langle M^k \alpha_{j_p}^{(k)}, \alpha_{j_p}^{(k)} \rangle_k}{\langle \alpha_{j_p}^{(k)}, \alpha_{j_p}^{(k)} \rangle_k}
$$

and the off-diagonal elements are the contribution from the aliasing vectors.

By Gershgorin's theorem, any eigenvalue λ of $\mathcal{M}_i$ must satisfy

$$
| \lambda - \langle M^k \alpha_n^{(k)}, \alpha_n^{(k)} \rangle_k | \leq \sum_{\substack{j\sim n\ (1) \\ j\neq n}} | \langle M^k \alpha_n^{(k)}, \alpha_j^{(k)} \rangle_k | \tag{5.7}
$$

for some $n \sim i\ (1)$. Therefore a bound on the asymptotic convergence rate, ρ , is given by

$$
\begin{aligned}
\rho &\leq \max_{|i|<N_1} \left(\max_{n\sim i\ (1)} \sum_{j\sim n\ (1)} | \langle M^k \alpha_n^{(k)}, \alpha_j^{(k)} \rangle_k | \right) \\
&= \max_{|i|<N_k} \sum_{j\sim i\ (1)} | \langle M^k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k |
\end{aligned} \tag{5.8}
$$

for the k-grid problem with meshsize $h_k = 1/N_k$ on the fine grid.

In section 5.3 we derive formulas for a bound on the righthand side of (5.8).

5.3 Derivation of Bounds on the Convergence Rate

For a fixed fine meshsize h, a given number of grids k, r smoothings and a damped Jacobi parameter ω, we derive formulas for a constant $C_{k,r,h,\omega} < 1$, independent of i which is a bound on the asymptotic convergence rate. In Section 5.4 we give values of these constants for various values of h, k and r using a typical value of ω.

By (5.8) it is enough to bound $\sum\limits_{j\sim i\ (1)} |\langle M^k\alpha_i^{(k)}, \alpha_j^{(k)}\rangle_k|$ independent of i. Divide the sum into two parts,

$$\begin{aligned}\sum_{j\sim i\ (1)} |\langle M^k\alpha_i^{(k)}, \alpha_j^{(k)}\rangle_k| &= |\langle M^k\alpha_i^{(k)}, \alpha_i^{(k)}\rangle_k| \\ &\quad + \sum_{\substack{j\sim i\ (0)\\ j\neq i}} |\langle M^k\alpha_i^{(k)}, \alpha_j^{(k)}\rangle_k| \\ &=: D_i + J_i \,,\end{aligned} \tag{5.9}$$

where D_i is the "diagonal part" and J_i is the "aliasing part" of the sum.

Let $i = (i_1, i_2)$, k, r, h and ω be fixed. Define

$$\xi_p = \xi_i^{(p)} = \cos^2\left(\frac{i_1\pi h_p}{2}\right), \tag{5.10a}$$

$$\eta_p = n_i^{(p)} = \cos^2\left(\frac{i_2\pi h_p}{2}\right), \tag{5.10b}$$

$$g_p = \langle G_p^r\alpha_i^{(p)}, \alpha_i^{(p)}\rangle_p \,, \tag{5.10c}$$

$$e_p = \langle\left(\sum_{\sigma=0}^{r-1} G_p^\sigma\right)\alpha_i^{(p)}, \alpha_i^{(p)}\rangle_p \tag{5.10d}$$

and

$$\nu_p = \nu_i^{(p)} \,, \tag{5.10e}$$

where the i, r, h and ω dependence has been suppressed in the notation and only the grid level is displayed.

We have the following theorem.

Theorem 5.1

$$D_i = g_k - 2\omega c_k \nu_k \sum_{p=2}^{k-1} f_p \Big(\prod_{m=p+1}^{k} 4 g_m \xi_m^2 \eta_m^2 \Big) - \frac{c_k \nu_k}{c_1 \nu_1} \Big(\prod_{m=2}^{k} 4 g_m \xi_m^2 \eta_m^2 \Big), \tag{5.11a}$$

and

$$J_i \leq 2\omega c_k \nu_k \sum_{p=2}^{k-1} f_p \Big(1 - (\prod_{m=p+1}^{k} \xi_m \eta_m) \Big) \Big(\prod_{m=p+1}^{k} 4 \mid g_m \mid \xi_m \eta_m \Big) \tag{5.11b}$$

$$+ \frac{c_k \nu_k}{c_1 \nu_1} \Big(1 - (\prod_{m=2}^{k} \xi_m \eta_m) \Big) \Big(\prod_{m=2}^{k} 4 \mid g_m \mid \xi_m \eta_m \Big) \ .$$

Remark 5.3 Theorem 5.1 allows us to obtain a bound on the asymptotic convergence rate that is no more complicated than the diagonal elements themselves.

Before proving Theorem 5.1 we find expressions for the inner products

$$\langle M^k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k.$$

Lemma 5.1

For any $j \sim i\ (1)$,

$$\langle M^k \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k = g_k \langle \alpha_i^{(k)}, \alpha_j^{(k)} \rangle_k - 2\omega c_k \nu_k \sum_{p=2}^{k-1} f_p \Big(\prod_{m=p+1}^{k} 4 g_m \xi_m \eta_m \Big) \langle \alpha_i^{(p)}, I_k^p \alpha_j^{(k)} \rangle_p$$

$$- \frac{c_k \nu_k}{c_1 \nu_1} \Big(\prod_{m=1}^{k} 4 g_m \xi_m \eta_m \Big) \langle \alpha_i^{(1)}, I_k^1 \alpha_j^{(k)} \rangle_1 \ . \tag{5.12a}$$

Proof of Lemma 5.1

We prove by induction that for every $s \le k$,

$$\langle M^s \alpha_i^{(s)}, \alpha_j^{(s)} \rangle_s = g_s \langle \alpha_i^{(s)}, \alpha_j^{(s)} \rangle_s - 2\omega c_s \nu_s \sum_{p=2}^{s-1} f_p \Big(\prod_{m=p+1}^{s} 4 g_m \xi_m \eta_m \Big) \langle \alpha_i^{(p)}, I_s^p \alpha_j^{(s)} \rangle_p$$
$$- \frac{c_s \nu_s}{c_1 \nu_1} \Big(\prod_{m=1}^{s} 4 g_m \xi_m \eta_m \Big) \langle \alpha_i^{(1)}, I_s^1 \alpha_j^{(s)} \rangle_1 . \quad (5.12b)$$

Taking $s = k$ gives (5.12a).

We start with $s = 2$. From (5.3), (5.10) and (2.12),

$$\begin{aligned} \langle M^2 \alpha_i^{(2)}, \alpha_j^{(2)} \rangle_2 &= \langle G_2^r \alpha_i^{(1)}, \alpha_j^{(2)} \rangle_2 - \langle A_1^{-1} I_2^1 G_2^r A_2 \alpha_i^{(2)}, I_2^1 \alpha_j^{(2)} \rangle_1 \\ &= g_2 \langle \alpha_i^{(2)}, \alpha_j^{(2)} \rangle_2 - \frac{\nu_2}{\nu_1} g_2 \xi_2 \eta_2 \langle \alpha_i^{(1)}, I_2^1 \alpha_j^{(2)} \rangle_1 . \end{aligned} \quad (5.13)$$

Using $4c_2 = c_1$ gives us (5.12a) for $k = 2$.

Assume (5.12a) is true for $k = s-1$ grids, $s \ge 3$. For the s-grid problem, (5.3), (5.10) and (2.12) give

$$\begin{aligned} \langle M^s \alpha_i^{(s)}, \alpha_j^{(s)} \rangle_s &= \langle G_s^r \alpha_i^{(s)}, \alpha_j^{(s)} \rangle_s - \langle (I - M^{s-1}) A_{s-1}^{-1} I_s^{s-1} G_s^r A_s \alpha_i^{(s)}, I_s^{s-1} \alpha_j^{(s)} \rangle_{s-1} \\ &= g_s - \frac{\nu_s}{\nu_{s-1}} \xi_s \eta_s g_s \langle \alpha_i^{(s-1)}, I_s^{s-1} \alpha_j^{(s)} \rangle_{s-1} \\ &\quad + \frac{\nu_s}{\nu_{s-1}} \xi_s \eta_s g_s \langle M^{s-1} \alpha_i^{(s-1)}, I_s^{s-1} \alpha_j^{(s)} \rangle_{s-1} . \end{aligned} \quad (5.14)$$

We factor $1 - g_{s-1} = 2\omega c_{s-1} \nu_{s-1} f_{s-1}$. Using the inductive hypothesis and using $4c_s = c_{s-1}$ finishes the proof. ∎

Proof of Theorem 5.1

(a) Using Lemma 5.1 with $j = i$, (5.10c) and (2.12) proves (5.11a).

(b) To prove (5.11b), split the grid levels by partitioning the $j \sim i$ (1), $j \ne i$. See Figure 3.1 for a schematic illustration for $k = 3$. For each $n = 1, \dots, k-1$ consider the j's such that $j \sim i$ (n) but $j \not\sim i$ $(n+1)$. Lemma 5.1, Lemma 2.1 and Lemma

2.3 show that

$$J_i = \sum_{n=1}^{k-1} \sum_{\substack{j\sim i\ (n)\\ j\not\sim i\ (n+1)}} |\langle M^k \alpha_i^{(k)}, \alpha_j^{(k)}\rangle_k | \tag{5.15}$$

$$\leq 2\omega c_k \nu_k \sum_{n=1}^{k-1}(1-\xi_{n+1}\eta_{n+1}) \sum_{p=1}^{n} f_p \Big(\prod_{m=p+1}^{n} \xi_m\eta_m \Big)\Big(\prod_{m=p+1}^{k} 4 \mid g_m \mid \xi_m\eta_m \Big)$$

$$+\frac{c_k\nu_k}{c_1\nu_1}\Big[\sum_{n=1}^{k-1}(1-\xi_{n+1}\eta_{n+1})\Big(\prod_{m=1}^{n}\xi_m\eta_n\Big)\Big]\Big(\prod_{m=1}^{k} 4 \mid g_m \mid \xi_m\eta_m\Big).$$

Changing the order of summation gives

$$J_i \leq 2\omega c_k\nu_k \sum_{p=2}^{k-1} f_p \Big[\sum_{n=p}^{k-1}(1-\xi_{n+1})\ \Big(\prod_{m=p+1}^{n} \xi_m\eta_m\Big)\Big]\Big(\prod_{m=p+1}^{k} 4 \mid g_m \mid \xi_m\eta_m\Big) \tag{5.16}$$

$$+\frac{c_k\nu_k}{c_1\nu_1}\ \Big[\sum_{n=1}^{k-1}(1-\xi_{n+1}\eta_{n+1})\Big(\prod_{m=p+1}^{n} \xi_m\eta_m\Big)\Big]\Big(\prod_{m=1}^{k} 4 \mid g_m \mid \xi_m\eta_m\Big).$$

The quantitites in the square brackets in (5.16) equal

$$1-\prod_{m=p+1}^{k} \xi_m\eta_m$$

and

$$1-\prod_{m=1}^{k} \xi_m\eta_m$$

respectively, and therefore (5.11b) has been proved. ∎

We use this theorem to find bounds on the multigrid V-cycle asymptotic convergence rate for the k-grid problem with a given damped Jacobi parameter ω and r iterations per smooth. The results are given in the next section.

5.4 Computed values of the asymptotic convergence bounds

Ideally, one would be able to compute k-grid convergence bounds independent of h . The $4^{k-1} \times 4^{k-1}$ matrix, $\mathcal{M}_i$, can be written as a $4^{k-1} \times 4^{k-1}$ matrix, $\mathcal{M}(\xi,\eta)$, with variable entries depending on the continuous variables ξ and $\eta \in (0,1)$ evaluated at $\xi = \xi_i^{(k)}$ and $\eta = \eta_i^{(k)}$.

In the two grid case one could get an analytic formula for the characteristic equation of $\mathcal{M}(\xi, \eta)$ (a polynomial of degree 4 for fixed ξ, η), find analytic expressions for the eigenvalues and then find the supremum of these expressions over all ξ and η in the unit square. This would give an exact 2-grid asymptotic convergence rate independent of h . In practice this is too much work even in the simple 2-grid case. Instead, one chooses a value of $h = 1/N$ and computes the spectral radii of $\mathcal{M}_i$ for each i, $\mid i \mid< N$, keeping track of the largest. One then repeats the procedure for different values of h and so constructs a table as in [12] see Table 5.1. From such tables one can predict the h-independent convergence rates.

In the k grid problem, $k > 2$ each $\mathcal{M}_i$ is a $4^{k-1} \times 4^{k-1}$ matrix and therefore computing the spectial radius for each i, $\mid i \mid < N$ is expensive, especially for small h . We therefore use Theorem 5.1 and Gershgorin's Theorem to compute a bound on the spectral radius of $\mathcal{M}_i$ for each i . This amounts to roughly twice the work of just evaluating the diagonal elements.

The sharpest bounds on the asymptotic convergence rates for the analysis of the V-cycle are obtained by these techniques when no smoothing is performed on the coarse-to-fine part of the cycle, i.e., $s = 0$ in step d. This is called an $M\backslash$ cycle. The symmetric cycle, i.e., $s = r$, is called an MG cycle. We consider two discretizations of the Laplacian, the five point discretization, $B_p = A_p$, as given in Section 2, and a certain nine point discretization given by the following stencil:

$$\tilde{A}_p = \frac{1}{3h_p^2} \begin{bmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{bmatrix}_{h_p} . \tag{5.17}$$

The corresponding V-cycles will be denoted by, e.g., $M_5\backslash$, or MG_9, to indicated which discretization is being used.

We consider a $M_5\backslash$ algorithm and compare our theoretical bounds to the experimentally observed asymptotic convergence rates. In order to compare our two grid bounds to the exact two grid convergence rates obtained by the model problem analysis in [8], we consider a damped Jacobi paramete $\omega = 4/5$. Experimentally, this is a good choice, though its optimality depends on the number of smoothings and the number of grids. We take $r = 1$, 2, 3 or 4 smoothings (smoothing only from fine to coarse meshes). Tables 5.2-5.5 show the convergence bounds for commonly used meshsizes. Table 5.6 indicates the limiting behaviour of these rates for very small h and large number of grids. The experimentally observed asymptotic convergence rates are shown in Table 5.7 for $r = 1, 2, 3, 4$, $\omega = 4/5$ and $h = 1/64$. For exact two grid convergence rates, see Table 5.1

In practice, as k increases there is not as much degradation in the convergence rate as Tables 5.1-5.7 would indicate.

We compare our bounds to the finite element bounds of [8], using the MG_9 cycle given by taking $B_p = \tilde{A}_p$ and $s = r$. The comparison is possible because the operators $\tilde{A}_p$ satisfy:

$$\tilde{A}_{p-1} = I_p^{p-1} \tilde{A}_p I_{p-1}^p \quad \text{for } p = 1, 2, \ldots, k. \tag{5.18}$$

Eigenvectors of A_p are also eigenvectors of $\tilde{A}_p$. We also note that for a symmetric V-cycle, convergence bounds in the energy norm are equivalent to asymptotic convergence bounds given by the spectral radius. Our bounds are given in Table 5.8 for $\omega = 3/4$, $h = 1/64$, and $r = 1, 2, 3, 4$. In the next to the last column of Table 5.8 we show the bounds (which are independent of the number of grids used) obtained by the methods of [8]. We also calculate the exact two grid convergence rates for MG_9, as in [12]. These numbers are given in the last column of Table 5.8. In this symmetric case, at least for small r, our bounds are larger than the finite element bounds because in the Fourier analysis we essentially throw away the post smoothing factors in the off-diagonal terms in order to be able to apply Lemma 5.1.

Table 5.1 $M_5\backslash$ Two grid asymptotic convergence rates $\omega = .8$

h	$r=1$	$r=2$	$r=3$	$r=4$
1/16	.592	.351	.208	.135
1/32	.598	.358	.214	.137
1/64	.600	.359	.216	.137
1/128	.600	.360	.216	.137

Table 5.2 $M_5\backslash$ Asymptotic convergence bounds $\omega = .8\ ,\ r = 1$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.615	.719	.715				
1/32	.622	.749	.769	.750			
1/64	.624	.758	.797	.800	.787		
1/128	.625	.760	.808	.826	.820	.815	
1/256	.625	.761	.812	.835	.835	.830	.828

Table 5.3 $M_5\backslash$ Asymptotic convergence bounds $\omega = .8\ ,\ r = 2$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids	8 grids
1/16	.369	.454	.455				
1/32	.370	.460	.481	.481			
1/64	.370	.466	.490	.491	.491		
1/128	.370	.467	.495	.499	.500	.499	
1/256	.370	.468	.495	.502	.505	.505	.504

Table 5.4 $M_5\backslash$ Asymptotic convergence bounds $\omega = .8\ ,\ r = 3$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids
1/16	.274	.348	.367			
1/32	.274	.348	.367	.372		
1/64	.275	.350	.370	.372	.373	
1/128	.275	.350	.371	.376	.376	.376

Table 5.5 $M_5\backslash$ Asymptotic convergence bounds $\omega = .8$, $r = 4$

h	2 grids	3 grids	4 grids	5 grids	6 grids	7 grids
1/16	.220	.284	.302			
1/32	.221	.284	.302	.307		
1/64	.221	.284	.302	.307	.308	
1/128	.221	.284	.302	.307	.308	.309

Table 5.6 $M_5\backslash$ Asymptotic convergence bounds for small h

$\omega = .8$

h	$r = 1$	$r = 2$	$r = 3$	$r = 4$
1/2048 11 grids	.843	.5105	.37779	.3087905
1/4096 12 grids	.846	.5111	.37777	.3087916

Table 5.7 $M_5\backslash$ Experimental asymptotic convergence rates

$$\omega = .8, \quad h = 1/64$$

r	2 grids	3 grids	4 grids	5 grids	6 grids
1	.600	.600	.600	.600	.600
2	.360	.360	.360	.360	.360
3	.216	.228	.233	.242	.246
4	.137	.158	.171	.181	.193

Table 5.8 MG_9 A comparison of the theoretical bounds

$$\omega = .75, \quad h = 1/64$$

r	2 grids	3 grids	4 grids	5 grids	6 grids	bounds from [8]	exact 2 grid conv. rates
1	.686	.717	.816	.860	.879	.40	.249
2	.275	.299	.348	.362	.364	.25	.067
3	.121	.147	.161	.162	.162	.18	.040
4	.079	.114	.124	.124	.124	.14	.029

APPENDIX

A.1 Proof of Theorem 3.4

Fix $i = (i_1, i_2), h, k, r$ and ω as in Section 3.3. Define ξ_p, η_p, g_p, e_p and ν_p as is (3.16a-e). As seen in the proof of Lemma 3.1,

$$D_i^{(p)} := \langle M_p \alpha_i^{(p)}, \alpha_i^{(p)} \rangle_p = \langle (I - G_p^{2r})\, A_p^{-1} \alpha_i^{(p)}, \alpha_i^{(p)} \rangle_p + \langle M_{p-1} I_p^{p-1} G_p^r \alpha_i^{(p)},\, I_p^{p-1} G_p^r \alpha_i^{(p)} \rangle_{p-1}. \tag{A.1}$$

Therefore a recursion formula for $D_i^{(p)}$ is

$$D_i^{(p)} = a_p + b_p D_i^{(p-1)} \qquad p = 1, 2, \ldots, k \tag{A.2}$$

where a_p and b_p are given by:

$$a_p = 2\omega c_p e_p \qquad p = 1, \ldots, k \tag{A.3a}$$

$$b_p = g_p^2 \xi_p^2 \eta_p^2 \langle \alpha_i^{(p-1)}, \alpha_i^{(p-1)} \rangle_{p-1} \qquad p = 1, \ldots, k \tag{A.3b}$$

$$b_0 = 0. \tag{A.3c}$$

The following four lemmas are all proved by direct calculation.

Lemma A.1

For each $p = 1, 2, \ldots, k$

$$\text{a.)} \quad a_p \le 4r\omega c_p \tag{A.4a}$$

$$\text{and} \quad \text{b.)} \quad a_p \ge 4\omega(1-\omega)c_p. \tag{A.4b}$$

Lemma A.2

For each $p = 2, \ldots, k$

$$\text{a.)} \quad b_p \le 1 \tag{A.5a}$$

and if $c_p\nu_p \le 1/4$,

$$\text{b.)} \quad b_p \ge (1 - 4(1 + r\omega)c_p\nu_p)\,. \tag{A.5b}$$

Lemma A.3

For each $p = 2, \ldots, k$

$$\text{a.)} \quad \nu_p/\nu_{p-1} \le \frac{1}{\xi_p\eta_p} \tag{A.6a}$$

$$\text{and} \quad \text{b.)} \quad \nu_p/\nu_{p-1} \ge 1. \tag{A.6b}$$

Lemma A.4

If $\dfrac{\beta}{4^{\alpha+1}} \le c_p\nu_p \le \dfrac{\beta}{4^\alpha}$, $p = 2, \ldots, k$ and $\dfrac{1}{2} < \beta < 2$, then

$$\text{a.)} \quad c_{p-n}\nu_{p-n} \le \frac{\beta}{4^{\alpha-n}} \tag{A.7a}$$

$$\text{and} \quad \text{b.)} \quad c_{p-n}\nu_{p-n} \ge \frac{\beta}{4^{\alpha-n+1}}\left(1 - \frac{2}{3}\frac{\beta}{4^{\alpha-n+1}}\right). \tag{A.7b}$$

Proof of Lemma A.1

Inequality (A.4a) follows immediately from the inequality

$$1 - (1 - x)^{2r} \le 2rx \tag{A.8}$$

since $|1 - x| \le 1$ where $x = 2\omega c_p\nu_p$.

Using the inequality

$$1 - (1 - x)^{2r} \ge x(2 - x) \qquad \text{for all } x \text{ such that } |1 - x| \le 1, \tag{A.9}$$

it is clear that

$$a_p \ge 2\omega c_p\,(2 - 2\omega c_p\nu_p)\,, \tag{A.10}$$

from which follows (A.4b) since $0 < c_p\nu_p < 1$. ∎

Proof of Lemma A.2

Since

$$c_p\nu_p = \frac{2-\xi_p-\eta_p}{2}, \tag{A.11}$$

where $0 < \xi_p, \eta_p < 1$ and $|g_p| = |1-2\omega c_p\nu_p| < 1$, (A.5a) is obvious.

If $c_p\nu_p \leq 1/4$, then $(1-\xi_p)$ and $(1-\eta_p) < 1/2$ and therefore

$$\langle \alpha_i^{(p-1)},\, \alpha_i^{(p-1)} \rangle_{p-1} = 1.$$

Moreover,

$$\xi_p^2\eta_p^2 \geq 1 - 4c_p\nu_p. \tag{A.12}$$

It is also clear that $(1-2\omega c_p\nu_p)^{2r} \geq 1 - 4r\omega c_p\nu_p$. Combining these inequalities gives

$$\begin{aligned} b_p &\geq (1-4r\omega c_p\nu_p)(1-4c_p\nu_p) \\ &\geq 1 - 4(1+r\omega)c_p\nu_p. \quad \blacksquare \end{aligned} \tag{A.13}$$

Proof of Lemma A.3

Since $0 < \xi_p, \eta_p < 1$,

$$-\xi_p(1-\xi_p)(1-\eta_p) - \eta_p(1-\xi_p)(1-\eta_p) \leq 0. \tag{A.14}$$

Factoring the lefthand side gives

$$\xi_p\eta_p(2-\xi_p-\eta_p) \leq \xi_p(1-\xi_p) + \eta_p(1-\eta_p). \tag{A.15}$$

Recall that

$$\nu_p = \frac{4(2-\xi_p-\eta_p)}{h_{p-1}^2}$$

and

$$\nu_{p-1} = \frac{4(2-\xi_{p-1}-\eta_{p-1})}{h_{p-1}^2} = \frac{4(\xi_p(1-\xi_p) + \eta_p(1-\eta_p))}{h_p^2}. \tag{A.16}$$

Thus by (A.15)

$$\frac{\nu_p}{\nu_{p-1}} \leq \frac{1}{\xi_p\eta_p}.$$

The second inequality, (A.6b), is clear since

$$\frac{\nu_p}{\nu_{p-1}} = \frac{(1-\xi_p) + (1-\eta_p)}{\xi_p(1-\xi_p) + \eta_p(1-\eta_p)} > 1 \tag{A.17}$$

and $0 < \xi_p, \eta_p < 1$. $\blacksquare$

Proof of Lemma A.4

If

$$\gamma \le c_p\nu_p \le \tau \le 1/4 \tag{A.18a}$$

then

$$4\gamma(1-2\gamma) \le c_{p-1}\nu_{p-1} \le 4\tau(1-2\tau). \tag{A.18b}$$

Note that this is just the calculus problem: Find the maximum and minimum of $f(x,y) = x(1-x)+y(1-y)$ in $\bar{\Omega} = \{(x,y)\colon\ 2\gamma \le x+y \le 2\tau,\ x \ge 0,\ y \ge 0\}$, and the solution is straightforward.

By induction, it is easy to see that

$$\frac{\beta}{4^{\alpha-n+1}} \prod_{j=0}^{n-1} \left(1 - \frac{2\beta}{4^{\alpha+1-j}}\right) \le c_{p-n}\nu_{p-n} \le \frac{\beta}{4^{\alpha-n}}. \tag{A.19}$$

By (A.18) this is true for $n = 1$, i.e.

$$\frac{\beta}{4^\alpha}(1 - \frac{2\beta}{4^{\alpha+1}}) \le c_{p-1}\nu_{p-1} \le \frac{\beta}{4^{\alpha-1}}(1 - \frac{2\beta}{4^\alpha}) \le \frac{\beta}{4^{\alpha-1}}. \tag{A.20}$$

Assume (A.19) is true for $c_{p-n+1}u_{p-n+1}$, then

$$\begin{aligned} c_{p-n}\nu_{p-n} &\ge 4\frac{\beta}{4^{\alpha-n+2}} \prod_{j=0}^{n-2} \left(1 - \frac{2\beta}{4^{\alpha+1-j}}\right) \left[1 - \frac{2\beta}{4^{\alpha-n+2}} \prod_{j=0}^{n-2} \left(1 - \frac{2\beta}{4^{\alpha+1-j}}\right)\right] \\ &\ge \frac{\beta}{4^{\alpha-n+1}} \left(\prod_{j=0}^{n-2} \left(1 - \frac{2\beta}{4^{\alpha+1-j}}\right)\right) \left(1 - \frac{2\beta}{4^{\alpha-n+2}}\right) \qquad \text{(A.21)} \\ &= \frac{\beta}{4^{\alpha-n+1}} \prod_{j=0}^{n-1} \left(1 - \frac{2\beta}{4^{\alpha+1-j}}\right). \end{aligned}$$

Using $\prod_{j=0}^{n} (1-x_j) \ge 1 - \sum_{j=0}^{n} x_j$ gives

$$\begin{aligned} \prod_{j=0}^{n-1} \left(1 - \frac{2}{4^{\alpha+1-j}}\right) &\ge 1 - \frac{2}{4^{\alpha-n+2}} \left(\sum_{j=0}^{n-1} \frac{1}{4^j}\right) \\ &\ge 1 - \frac{2}{3} \cdot \frac{1}{4^{\alpha-n+1}}. \end{aligned} \tag{A.22}$$

The upper bound for $c_{p-n}\nu_{p-n}$ is obvious. ∎

Proof of Theorem 3.4 part a.

a.) $$\frac{\omega(1-\omega)}{\max(2, d(1+r\omega))} h_1^2 \le D_i^{(k)} \le \frac{2r\omega}{3} h_1^2, \qquad \text{for } \nu_k \le \frac{d}{h_1^2}.$$

From lemmas A.1a and A.2a, $p = k, k-1, \dots, 2$,

$$D_i^{(k)} \le 4r\omega \left(\sum_{p=2}^{k} c_p \right) + D_i^{(1)}. \tag{A.23}$$

On the coarsest grid, $D_i^{(1)} \le 4r\omega\, c_1$. Hence by the definition of the c_p's (2.15)

$$D_i^{(k)} \le 4r\omega \left(\sum_{p=1}^{k} c_p \right) \le \frac{16r\omega}{3} c_1. \tag{A.24}$$

Using $c_1 = h_1^2/8$ gives the upper bound

$$D_i^{(k)} \le \frac{2}{3} r\omega h_1^2. \tag{A.25}$$

To get the lower bound, use an induction argument. By lemma A.16,

$$D_i^{(1)} \ge 4\omega(1-\omega)\, c_1 = \frac{\omega(1-\omega)h_1^2}{2}. \tag{A.26}$$

Let $1 \le p < k$ and assume that

$$D_i^{(p)} \ge \frac{\omega(1-\omega)h_1^2}{\max(2,\, d(1+r\omega))}. \tag{A.27}$$

By rearranging terms, using lemmas A.1b and A.2b (which can be used since $\nu_k < d/h_1^2$ implies $c_p\nu_p \le d/(8\cdot 4^p) \le 1/4$ by lemma A.4a) it is seen that:

$$D_i^{(p+1)} \ge \frac{\omega(1-\omega)h_1^2}{\max(2,\, d(1+r\omega))} + 4\omega(1-\omega)c_{p+1}\left(1 - \frac{(1+r\omega)\nu_{p+1}\, h_1^2}{\max(2,\, d(1+r\omega))}\right). \tag{A.28}$$

Lemma A.3b guarantees that $\nu_p \le \nu_k \le \dfrac{d}{h_1^2}$ and therefore the last term in (A.28) is positive and can be thrown out. This proves part a.) of Theorem 3.4. ∎

Proof of Theorem 3.4 part b.

b.) $$\frac{d\omega(1-\omega)}{8(1+r\omega)\nu_k} \leq D_i^{(k)} \leq \frac{1}{\nu_k} \quad \text{for } \nu_i^{(k)} > \frac{d}{h_1^2}.$$

Using the definition of c_k, $\nu_k > d/h_1^2$ implies $c_k \nu_k \geq d/(2 \cdot 4^{k+1})$. For each p_1, $\lambda_i^{(p)} \leq 1$. To see this, first note that

$$\lambda_i^{(1)} = 1 - (1 - 2\omega c_1 \nu_1)^{2r} \leq 1. \tag{A.29}$$

Lemma A.3a together with the definition of b_p imply

$$b_p \leq \frac{(1 - 2\omega c_p \nu_p)^{2r} \nu_{p-1}}{\nu_p}. \tag{A.30}$$

Combining (A.29) and (A.30) with the definition of a_p, gives $D_i^{(p)} \leq 1/\nu_p$.

For the lower bound, divide the argument into two cases. Define

$$\gamma = [\log_4 2(1 + r\omega)] \tag{A.31}$$

where $[x]$ means the greatest integer in x.

$$\textit{case 1} \qquad c_k \nu_k \geq \frac{d}{2 \cdot r^{\gamma}} \tag{A.32}$$

$$\textit{case 2} \qquad \frac{d}{2 \cdot 4^{\alpha+1}} \leq c_k \nu_k \leq 1 \quad \text{for some integer } \alpha,\ \gamma \leq \alpha \leq k. \tag{A.33}$$

By the definition of γ,

$$\frac{1}{2}(1 + r\omega) \leq 4^{\gamma} \leq 2(1 + r\omega). \tag{A.34}$$

For *case 1*, Lemma A.1b gives

$$a_k \geq \frac{4\omega(1-\omega)}{\nu_k} \frac{d}{2 \cdot 4^{\gamma}} \geq \frac{d\omega(1-\omega)}{(1+r\omega)\nu_k}. \tag{A.35}$$

For *case 2*, look at the finest grid on which the eigenvector is "rough enough". For $\hat{p} = k - \alpha + \gamma$, lemma A.4 shows that

$$\frac{d}{2 \cdot 4^{\gamma+1}} \left(1 - \frac{d}{3 \cdot r^{\gamma+1}}\right) \leq c_{\hat{p}} \nu_{\hat{p}} \leq \frac{d}{2 \cdot 4^{\gamma}}. \tag{A.36}$$

Therefore on $\Omega^{\hat{p}}$,

$$D_i^{(\hat{p})} \geq a_{\hat{p}} \geq 4\omega(1-\omega)c_{\hat{p}} \geq \frac{d(1-\omega)\omega}{8(1+r\omega)\nu_{\hat{p}}}. \tag{A.37}$$

Now this information needs to get back to the fine grid, Ω^k. On Ω^p, for $p > \hat{p}$, lemma A.4 says

$$\frac{d}{2 \cdot 4^{\alpha-k+p+1}}\left(1 - \frac{d}{3 \cdot 4^{\alpha-k+p+1}}\right) \leq c_p\nu_p \leq \frac{d}{2 \cdot 4^{\alpha-k+p}}. \tag{A.38}$$

Now using lemmas A.16, A.26 and A.36 and rearranging terms,

$$\begin{aligned} D_i^{(p)} &\geq 4\omega(1-\omega)c_p + (1 - 4(1+r\omega)c_p\nu_p)\frac{\omega(1-\omega)}{8(1+r\omega)\nu_p} \\ &= \frac{d(1-\omega)\omega}{8(1+r\omega)\nu_p} + 4\omega(1-\omega)c_p\left[1 - \frac{d}{8}\right] \\ &\geq \frac{d(1-\omega)\omega}{8(1+r\omega)\nu_p}. \end{aligned} \tag{A.39}$$

Since this is true for all $p > \hat{p}$, take $p = k$. ∎

References

1. A. Bayliss, C. I. Goldstein and E. Turkel: "On Accuracy Conditions for the Numerical Computation of Waves", ICASE report #84-38 , 1984.

2. R. Chandra: "Conjugate Gradient Methods for Partial Differential Equations", Research Report #229, Department of Computer Science, Yale University, 1981.

3. C. D. Conte and C. deBoor: *Elementary Numerical Analysis, An Algorithmic Approach*, McGraw–Hill, Inc., 3rd Edition, 1980.

4. N. H. Decker: Ph.D. Thesis, Department of Mathematics, University of Wisconsin–Madison, in preparation.

5. V. Faber, T. A. Manteuffel and S. V. Parter: "On the Equivalence of Operators and the Implications to Preconditioned Iterative Methods for Elliptic Problems", to appear.

6. G. Forsythe and W. Wasow: *Finite–Difference Methods for Partial Differential Equations*, John Wiley and Sons, Inc., 1960.

7. C. I. Goldstein: "Preconditioned Iterative Methods Applied to Singularly Perturbed Elliptic Boundary Value Problems", Report #51915, Brookhaven National Labs, 1985.

8. J. Mandel, S. F. McCormick and R. Bank: Chapter 5, "Variational Multigrid Theory". In: S. F. McCormick, Editor, Multigrid Methods, SIAM Frontiers in Applied Math. 5 , to appear.

9. T. A. Manteuffel: "An Incomplete Factorization Technique for Positive Definite Linear Systems", Mathematics of Computation, Vol. 34, Number 150, April 1980, pp.473-479.

10. S. F. McCormick: Chapter 1, "Introduction". In: S. F. McCormick, Editor, Multigrid Methods, SIAM Frontiers in Applied Math., to appear.

11. A. H. Shatz and L. B. Wahlbin: "On the Finite Element Method for Singularly Perturbed Reaction–Diffusion Problems in Two and One Dimensions", Mathematics of Computation, Vol. 40, Number 161, pp.47-89, 1983.

12. K. Stüben and U. Trottenberg: "Multigrid Methods : Fundamental Algorithms, Model Problem Analysis and Applications", *Multigrid Methods*, Springer-Verlag Lecture Notes in Mathematics 960, 1981.