BIT Numerical Mathematics (2006) 46: 283–306
DOI: 10.1007/s10543-006-0052-5
Published online: 3 May 2006 – © Springer 2006

FOURIER ANALYSIS OF MULTIGRID FOR A MODEL TWO-DIMENSIONAL CONVECTION-DIFFUSION EQUATION*

HOWARD C. ELMAN[1,**] and ALISON RAMAGE[2]

[1] *Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA. email: elman@cs.umd.edu*

[2] *Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, Scotland. email: A.Ramage@strath.ac.uk*

Abstract.

We present a Fourier analysis of multigrid for the two-dimensional discrete convection-diffusion equation. For constant coefficient problems with grid-aligned flow and semi-periodic boundary conditions, we show that the two-grid iteration matrix can be reduced via a set of orthogonal transformations to a matrix containing individual 4×4 blocks. This enables a trivial computation of the norm of the iteration matrix demonstrating rapid convergence in the case of both small and large mesh Peclet numbers, where the streamline-diffusion discretisation is used in the latter case. We also demonstrate that these results are strongly correlated with the properties of the iteration matrix arising from Dirichlet boundary conditions.

AMS subject classification (2000): 65F10, 65N22, 65N30, 65N55.

Key words: multigrid, convection-diffusion, convergence.

1 Introduction.

In this paper, we present some analysis of a multigrid solution method for the linear convection-diffusion equation

$$\begin{aligned} -\epsilon\nabla^2 u(x,y) + \mathbf{w}\cdot\nabla u(x,y) &= f(x,y) \qquad \text{in } \Omega \\ u(x,y) &= g(x,y) \qquad \text{on } \delta\Omega \end{aligned} \tag{1.1}$$

where the divergence-free convective velocity field $\mathbf{w} = (w_1(x,y), w_2(x,y))$ and the small parameter ϵ are given. Multigrid is effective for this problem if suitable strategies (for example, streamline-diffusion discretisation [14] or problem-dependent grid transfer operators [27]) are used to handle instabilities associated with coarse grids, and if smoothers (such as flow-directed iterations [12]

* Received 2005. Accepted in revised form February 2006. Communicated by Anna Karin Tornberg.

** The work of this author was supported by the National Science Foundation under grant DMS0208015 and by the Department of Energy under grant DOEG0204ER25619.

or multidirectional sweeping strategies [20]) take into account properties of the flow. However, the state of analysis is less well-developed than for the standard 'model problem', the Poisson equation. Typical analyses, for example, Bank [1], Bramble, Pasciak and Xu [2], Mandel [16], Wang [25], generalise the results for the Poisson equation by treating the convection-diffusion operator as a perturbation of the Laplacian. Good surveys of this situation as well as other new results can be found in the recent papers of Reusken [21] and Olshanskii and Reusken [18]. Several results relevant to finite difference discretisations of convection-diffusion equations can be found in, for example, Stüben and Trottenberg [23] and Trottenberg et al. [24].

The work presented here is a Fourier analysis applied to the constant coefficient problem with grid aligned flow and semi-periodic boundary conditions. It shows that the two-grid iteration matrix can be transformed to a matrix containing 4×4 blocks, which in turn can be used to compute the convergence factor of the iteration matrix. Moreover, these convergence factors also essentially predict the behaviour of multigrid for the Dirichlet problem. This work generalises results for the one-dimensional problem given in Hackbusch [11]. It can be applied to any discretisation strategy that results in a computational molecule of a certain form (see (2.3)). It applies to both Galerkin discretisation in the case of small mesh Peclet numbers and streamline diffusion discretisation for large mesh Peclet numbers.

To fix ideas, we seek a solution u_h of the discrete weak form of equation (1.1),

$$\begin{aligned} &\epsilon(\nabla u_h, \nabla v) + (\mathbf{w}.\nabla u_h, v) + \sum \alpha^{el} (\mathbf{w} \cdot \nabla u_h, \mathbf{w} \cdot \nabla v)_{el} \\ &\qquad = (f_h, v) + \sum \alpha^{el} (f_h, \mathbf{w} \cdot \nabla v)_{el} \quad \forall v \in V_h, \end{aligned} \tag{1.2}$$

where h is a discretisation parameter, the test functions v are in a finite-dimensional subspace V_h of the Sobolev space $V = \mathcal{H}_0^1(\Omega)$, f_h is the $L^2(\Omega)$ orthogonal projection of f into V_h and the sums are taken over all elements in the discretisation. Choosing the test functions equal to a set of basis functions for V_h (bilinear functions on square elements here) leads to a sparse linear system whose solution can be used to recover the discrete solution u_h. The stabilisation parameters α^{el} are given by

$$\alpha^{el} = \frac{\delta^{el} h^{el}}{|\mathbf{w}|} \tag{1.3}$$

where h^{el} is a measure of element size, $|\mathbf{w}|$ represents the strength of the convective field within an element and $\delta^{el} \geq 0$ are parameters to be chosen. The choice of these parameters depends on the mesh Peclet number

$$P_h^{el} = \frac{h^{el}|\mathbf{w}|}{2\epsilon}.$$

Setting $\delta^{el} = 0$ on each element reduces (1.2) to the standard Galerkin discretisation: this is the usual practice when $P_h^{el} \leq 1$. If, however, $P_h^{el} > 1$, then the discrete solution obtained from the Galerkin method may exhibit non-physical oscillations. The effects of such oscillations may be minimised by choosing $\delta^{el} > 0$,

that is, by adding coercivity in the local flow direction. In this paper, we will adopt the strategy suggested by the analysis in [5], that is, we choose the stabilisation parameter to be

$$\delta^{el} = \max\left(0, \frac{1}{2} - \frac{\epsilon}{h^{el}}\right). \tag{1.4}$$

For full details of both Galerkin and streamline diffusion discretisations, see for example [8, 14, 15, 17, 19, 22].

Discretising (1.1) via (1.2) gives rise to a sparse linear system

$$A_f \mathbf{u} = \mathbf{f}. \tag{1.5}$$

In this paper, we are interested in analysing the convergence behaviour of multigrid methods for this system. For details of multigrid methods in general, see for example [3, 9, 26]. Here, we consider only a two-grid method, as multigrid convergence is closely related to the two-grid convergence (see for example [10]). Our two-grid method has the following matrix components:

- coefficient matrix A_f on a fine grid with grid parameter h_f;
- coefficient matrix A_c on a related coarse grid with grid parameter $h_c = 2h_f$, constructed by direct discretisation via (1.2);
- bilinear interpolation matrix P for prolongation from coarse to fine grid;
- restriction matrix P^T for restriction from fine to coarse grid;
- smoothing matrix S_A^ν corresponding to pre-smoothing with ν steps of line Gauss–Seidel iteration.

It is shown in [20] that, with stabilisation added as described above, this is an effective algorithm for two-dimensional convection-diffusion problems. For clarity of presentation, we have not included any post-smoothing steps; the analysis could be extended to include post-smoothing in an obvious way. A two-grid method with the components described above has the iteration matrix

$$M = \left(I - P A_c^{-1} P^T A_f\right) S_A^\nu, \tag{1.6}$$

where the contraction number is given by $\|M\|$ for any matrix norm.

The remainder of the paper is structured as follows. In the next section, we present details of the two-dimensional model problem studied, and the matrix transformations which we will use as a basis for our analysis. These transformations enable us to obtain a tractable expression for the 2-norm of the two-grid iteration matrix M by considering several individual problems with one-dimensional structure. In Sections 3 and 4, bounds are obtained for the norm of (1.6) corresponding to a semiperiodic version of the model problem for various values of mesh Peclet number. The usefulness of these bounds for predicting convergence behaviour for the full Dirichlet problem is then assessed in Section 5. Finally, Section 6 contains some concluding observations.

2 Matrix Transformation.

For the purposes of this analysis, we focus on the 'vertical wind' model problem

$$-\epsilon\nabla^2 u + \frac{\partial u}{\partial y} = 0 \quad \text{in } \Omega = (0,1)\times(0,1), \tag{2.1}$$

obtained by setting $\mathbf{w} = (0,1)$ and f=0 in (1.1). Dirichlet boundary conditions are applied on all boundaries, and we use a natural ordering of the unknowns on a uniform grid of square bilinear elements with $N_f = 1/h$ elements in each dimension. In what follows, we define orthogonal transformation matrices Q_f and Q_c and apply them as appropriate to the components of M in (1.6). We emphasise that these transformations can be applied to any discretisation technique whose stencil takes the form (2.3) (see below). We discuss the specific details for discretisation (1.2) in Section 5.

2.1 Coefficient matrices.

For model problem (2.1), discretisation (1.2) gives rise to a linear system (1.5) where the coefficient matrix A_f is of order n_f^2, with $n_f = N_f - 1$. The choice of stabilisation parameter δ is governed by the model problem mesh Peclet number

$$P_h = \frac{h}{2\epsilon}\,. \tag{2.2}$$

If $P_h < 1$, a Galerkin method is used ((1.2) with $\delta = 0$) and if $P_h \geq 1$, streamline diffusion is added ((1.2) with stabilisation parameter $\delta = \frac{1}{2} - \frac{\epsilon}{h}$). If the coefficients of the computational molecule are denoted by

$$\begin{array}{ccccc} m_4 & & m_3 & & m_4 \\ & \nwarrow & \uparrow & \nearrow & \\ m_2 & \leftarrow & m_1 & \rightarrow & m_2 \\ & \swarrow & \downarrow & \searrow & \\ m_6 & & m_5 & & m_6 \end{array}\,, \tag{2.3}$$

then the matrix A_f can be written as

$$A_f = \begin{bmatrix} M_1 & M_2 & & & 0 \\ M_3 & M_1 & M_2 & & \\ & \ddots & \ddots & \ddots & \\ & & M_3 & M_1 & M_2 \\ 0 & & & M_3 & M_1 \end{bmatrix} \tag{2.4}$$

where

$$M_1 = \text{tridiag}(m_2, m_1, m_2), \quad M_2 = \text{tridiag}(m_4, m_3, m_4),$$
$$M_3 = \text{tridiag}(m_6, m_5, m_6)$$

are all tridiagonal matrices of order n_f. Assume N_f is even. A related coarse grid of $N_c \times N_c$ elements is formed by removing every second node from the fine grid. This leads to the coefficient matrix A_c defined by direct discretisation via (1.2) on the $n_c \times n_c$ coarse grid (where $n_c = N_c - 1$). The stabilisation parameter for the coarse grid is chosen as above using the coarse grid mesh Peclet number

$$P_h^c = \frac{\bar{h}}{2\epsilon},$$

where the coarse grid spacing is $\bar{h} = 2h$.

Writing

$$\theta_j = \frac{j\pi}{N_f},$$

we have that the eigenvalues and eigenvectors of the blocks of A_f satisfy

$$\begin{aligned} M_1 \mathbf{v}_j^f &= \lambda_j^f \mathbf{v}_j \qquad \lambda_j^f = m_1 + 2m_2 \cos\theta_j \\ M_2 \mathbf{v}_j^f &= \sigma_j^f \mathbf{v}_j \qquad \sigma_j^f = m_3 + 2m_4 \cos\theta_j \\ M_3 \mathbf{v}_j^f &= \gamma_j^f \mathbf{v}_j \qquad \gamma_j^f = m_5 + 2m_6 \cos\theta_j \end{aligned} \tag{2.5}$$

for $j = 1, \ldots, n_f$, where the eigenvectors are

$$\mathbf{v}_j^f = \sqrt{\frac{2}{N_f}} \left[\sin\theta_j,\ \sin 2\theta_j,\ \ldots,\ \sin n_f\theta_j\right]^T. \tag{2.6}$$

We may therefore obtain the decomposition

$$T_f = (\mathcal{V}_f \Pi_f)^T A_f (\mathcal{V}_f \Pi_f) \equiv Q_f^T A_f Q_f \tag{2.7}$$

where

$$\mathcal{V}_f = I_f \otimes V_f, \qquad V_f = \left[\mathbf{v}_1^f, \mathbf{v}_2^f, \ldots, \mathbf{v}_{n_f}^f\right] \tag{2.8}$$

and the symbol $\otimes$ represents the Kronecker product (see e.g. [13, p. 243]). Here V_f represents the eigenvector transformation matrix, I_f is the $n_f \times n_f$ identity matrix, and Π_f is a permutation matrix of order n_f^2 which reorders the unknowns in vertical rather than horizontal lines; see [5], [6] for further details.

As the coarse grid coefficient matrix A_c is constructed in an exactly analogous way to A_f, we may identify coarse grid eigenvectors

$$\mathbf{v}_j^c = \sqrt{\frac{4}{N_f}} \left[\sin 2\theta_j,\ \sin 4\theta_j,\ \ldots,\ \sin 2n_c\theta_j\right]^T, \quad j = 1, \ldots, n_c \tag{2.9}$$

with corresponding eigenvalues λ^c, σ^c and γ^c (cf. (2.5)) and write

$$\mathcal{V}_c = I_c \otimes V_c, \quad V_c = \left[\mathbf{v}_1^c, \mathbf{v}_2^c, \ldots, \mathbf{v}_{n_c}^c\right]. \tag{2.10}$$

This leads to the transformation

$$T_c = Q_c^T A_c Q_c \tag{2.11}$$

where the matrix $Q_c = \mathcal{V}_c \Pi_c$ is orthogonal. The matrix T_c is block diagonal, with diagonal blocks $(T_c)_j = \text{tridiag}(\gamma_j^c, \lambda_j^c, \sigma_j^c)$, $j = 1, \ldots, n_c$, involving the coarse grid eigenvalues.

2.2 Smoothing Matrix.

Using the block matrix splitting

$$A_f = D_A - L_A - U_A,$$

where D_A is the block diagonal of A_f and L_A and U_A are the (strict) block lower and upper triangular parts, respectively, the line Gauss–Seidel smoothing matrix can be written as

$$S_A = (D_A - L_A)^{-1}U_A = I - (D_A - L_A)^{-1}A_f.$$

Hence

$$\begin{aligned} Q_f^T S_A Q_f &= I - \left(Q_f^T(D_A - L_A)Q_f\right)^{-1}\left(Q_f^T A_f Q_f\right) \\ &= I - (D_T - L_T)^{-1}T_f \equiv S_T \end{aligned} \tag{2.12}$$

where D_T and L_T are the diagonal and (strict) lower triangular parts of T_f, and S_T is block diagonal with blocks

$$(S_T)_j = I - (D_{T_j} - L_{T_j})^{-1}(T_f)_j, \quad j = 1, \ldots, n_f.$$

2.3 Prolongation and restriction matrices.

The prolongation matrix P (of dimension $n_f^2 \times n_c^2$) corresponding to bilinear interpolation from coarse to fine grid can be written as

$$P = L \otimes L = \begin{bmatrix} l_{11}L & \cdots & l_{1n_c}L \\ \vdots & \ddots & \vdots \\ l_{n_f 1}L & \cdots & l_{n_f n_c}L \end{bmatrix} \tag{2.13}$$

where L is the $n_f \times n_c$ one-dimensional linear interpolation matrix with entries

$$l_{ij} = \begin{cases} 1, & i = 2j \\ \frac{1}{2}, & i = 2j-1 \text{ or } i = 2j+1 \\ 0, & \text{otherwise} \end{cases}, \qquad i = 1, \ldots, n_f,\ j = 1, \ldots, n_c.$$

Using (2.6) and (2.9), it can be shown that the following relationships hold:

$$\begin{aligned} L^T \mathbf{v}_j^f &= \alpha_j \mathbf{v}_j^c & \alpha_j &= \tfrac{1}{\sqrt{2}}\left(1 + \cos\theta_j\right),\ j = 1, \ldots, n_c \\ L^T \mathbf{v}_{N_c}^f &= \mathbf{0}^f & &\text{(zero column vector of length } n_f) \\ L^T \mathbf{v}_{N_f - j}^f &= \alpha_{N_f - j}\mathbf{v}_j^c, & \alpha_{N_f - j} &= -\tfrac{1}{\sqrt{2}}\left(1 - \cos\theta_j\right),\ j = 1, \ldots, n_c. \end{aligned} \tag{2.14}$$

Introducing the $n_c \times n_f$ matrix $\mathcal{A}$ with entries

$$a_{jk} = \begin{cases} \alpha_j, & k = j \\ \alpha_k, & k = N_f - j \\ 0, & \text{otherwise} \end{cases},$$

we may write

$$L^T V_f = V_c \mathcal{A} \Rightarrow V_f^T L = \mathcal{A}^T V_c^T.$$

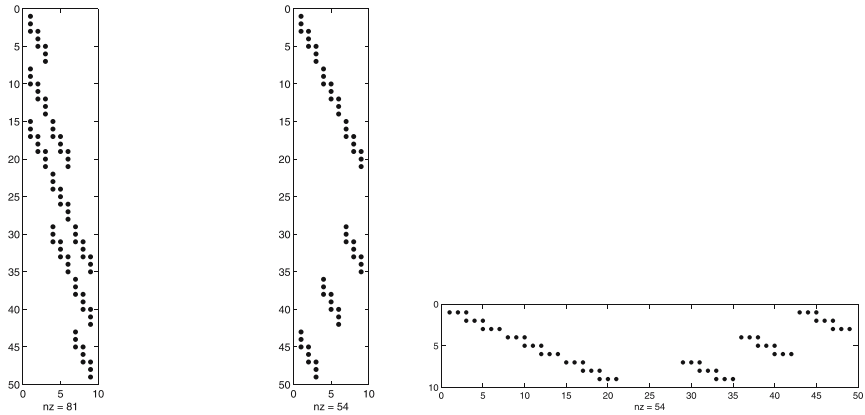
Combining this with (2.8), (2.10) and (2.13) gives a transformed prolongation matrix $\bar{P}$ as follows:

$$\begin{aligned} Q_f^T P Q_c &= \Pi_f^T \big(I_f \otimes V_f^T\big)(L \otimes L)(I_c \otimes V_c)\Pi_c = \Pi_f^T \big(L \otimes V_f^T L V_c\big)\Pi_c \\ &= \Pi_f^T (L \otimes \mathcal{A}^T)\Pi_c = \mathcal{A}^T \otimes L \equiv \bar{P}. \end{aligned} \tag{2.15}$$

Depictions of the sparsity patterns of P and $\bar{P}$ (for $N_f = 8$) are shown in Figure 2.1. Numbering the individual $n_f \times n_c$ blocks of $\bar{P}$ down from the top with $j = 1, \ldots, n_c$ and up from the bottom with $k = N_f - 1, \ldots, N_f - n_c$, we note that each block has the form

$$\bar{P}_j = \alpha_j L, \qquad \bar{P}_k = \alpha_k L$$

with α_j, α_k defined by (2.14).



(a) P. (b) $\bar{P}$. (c) $\bar{P}^T$.

Figure 2.1: Representative sparsity patterns of grid-transfer matrices, $N_f = 8$.

2.4 Iteration matrix.

It follows from (2.7), (2.11), (2.12) and (2.15) that the transformed version of the two-grid iteration matrix M in (1.6), which we will denote by $\bar{M}$, is

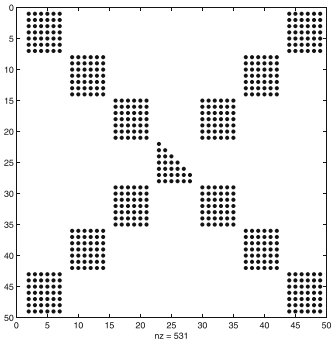
$$\begin{aligned} Q_f^T M Q_f &= Q_f^T \left[\big(I - P Q_c T_c^{-1} Q_c^T P^T Q_f T_f Q_f^T\big) S_A^\nu\right] Q_f \\ &= \big(I - \bar{P} T_c^{-1} \bar{P}^T T_f\big) Q_f^T \big(Q_f S_T Q_f^T\big)^\nu Q_f \\ &= \big(I - \bar{P} T_c^{-1} \bar{P}^T T_f\big)\, S_T^\nu \equiv \bar{M}. \end{aligned}$$

Recall that the matrices T_c^{-1}, T_f and S_T are all of block diagonal form. It can be seen from Figure 2.1 that the sparsity pattern of $\bar{P}$ has a block structure

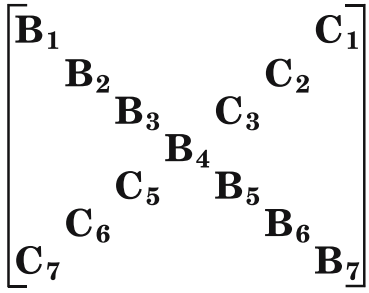
resembling an arrow pointing from left to right; similarly, the block structure of $\bar{P}^T$ resembles a V-shaped wedge. It is easily shown that (i) pre- or post-multiplying such "arrow" or "wedge" block matrices by block-diagonal matrices does not change the structure; and (ii) the product of an "arrow" block matrix and a "wedge" block matrix is a matrix whose nonzero blocks have the form of a cross; that is, the blocks of the product lie on the diagonal and counter-diagonal. The nonzero blocks in the matrix $\bar{M}$ therefore form this type of cross pattern: a sample sparsity plot for $N_f = 8$ is shown in Figure 2.2(a). We will denote the diagonal blocks of $\bar{M}$ by B_j, $j = 1, \ldots, n_f$ and the ante-diagonal blocks by C_j, $j = 1, \ldots, n_f$, noting that, with this notation, $B_{N_c} = C_{N_c}$. Applying a further permutation results in a block-diagonal matrix $\mathcal{M}$ (see Figure 2.2(c)). The diagonal blocks of $\mathcal{M}$ are given by

$$\mathcal{M}_j = \begin{bmatrix} B_j & C_j \\ C_k & B_k \end{bmatrix}, \quad j = 1, \ldots, n_c, \; k = N_f - j, \tag{2.16}$$

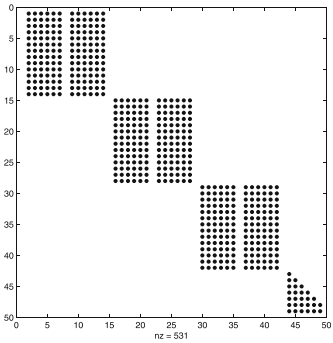
$$\mathcal{M}_{N_c} = B_{N_c},$$



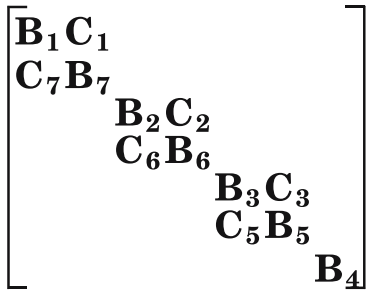
(a) $\bar{M}$.



(b) $\bar{M}$.



(c) $\mathcal{M}$.



(d) $\mathcal{M}$.

Figure 2.2: Representative sparsity patterns of iteration matrices, $N_f = 8$.

where the component matrices are defined by

$$\left.\begin{aligned} B_j &= \big(I - \bar{P}_j\big(T_c^{-1}\big)_j \bar{P}_j^T (T_f)_j\big) S_j^\nu \\ C_j &= -\bar{P}_j\big(T_c^{-1}\big)_j \bar{P}_k^T (T_f)_j S_k^\nu \\ B_k &= \big(I - \bar{P}_k\big(T_c^{-1}\big)_j \bar{P}_k^T (T_f)_k\big) S_k^\nu \\ C_k &= -\bar{P}_k\big(T_c^{-1}\big)_j \bar{P}_j^T (T_f)_j S_j^\nu \end{aligned}\right\} \quad j = 1, \ldots, n_c,\ k = N_f - j \tag{2.17}$$

$$B_{N_c} = S_{N_c}^\nu.$$

This notation is illustrated for $N_f = 8$ in Figures 2.2(b) and (d).

As all of the transformations which we have applied are orthogonal, we have that

$$\|M\|_2 = \|\bar{M}\|_2 = \|\mathcal{M}\|_2 = \max_{j=1,\ldots,N_c} \|\mathcal{M}_j\|_2. \tag{2.18}$$

We can therefore obtain the contraction number of the two-grid iteration for Dirichlet problem (2.1) by bounding the 2-norms of the blocks of $\mathcal{M}$.

3 The semiperiodic problem.

The result (2.18) is useful because it is cheaper to compute norms of the $2n_f \times 2n_f$ blocks of $\mathcal{M}$ than to compute the norm of the original $n_f^2 \times n_f^2$ iteration matrix M. We can gain further insight into multigrid convergence behaviour for the Dirichlet problem (2.1) by studying a closely related semiperiodic problem, namely (2.1) with the Dirichlet conditions on the boundaries $y = 0$, $y = 1$ replaced by the periodic condition $u(x, 0) = u(x, 1)$. For certain parameter regimes, analysis of this problem produces analytic formulae for the 2-norm in (2.18). In the next two sections, we describe the details of this analysis of the semiperiodic variant. Then, in Section 5, we state the extent to which these results are representative of the behaviour of the full Dirichlet problem (2.1).

For the semiperiodic problem, transformations analogous to (2.7) and (2.11) derived from exponential bases lead to block-diagonal matrices T_f^{per}, T_c^{per} in which the $n_f \times n_f$ matrices $(T_f)_j$, $(T_c)_j$ and S_j are replaced by $N_f \times N_f$ *circulant* matrices (recall $n_f = N_f - 1$) in the obvious way. For prolongation, each block $\bar{P}_j = \alpha_j L$ is extended to a $N_f \times N_c$ matrix by the addition of an extra row and column to L, with nonzero entries

$$l_{1,N_c} = \tfrac{1}{2}, \quad l_{n_f,N_c} = \tfrac{1}{2}, \quad l_{N_f,N_c} = 1.$$

We will denote these periodic variants with superscript *per*, and write, for example,

$$B_j^{per} = \big[I - \bar{P}_j^{per}\big(T_c^{per}\big)_j^{-1} (\bar{P}^{per})_j^T \big(T_f^{per}\big)_j\big] \big(S_j^{per}\big)^\nu \tag{3.1}$$

(cf. (2.17)).

We now define the vectors

$$\mathbf{w}_q^f = \sqrt{\frac{1}{N_f}}[e^{i(2\theta_q)}\ e^{i(4\theta_q)}\ \ldots\ e^{i(2N_f\theta_q)}]^T, \quad \theta_q = \frac{q\pi}{N_f},\ q = 1, \ldots, N_f \tag{3.2}$$

which are the orthonormal eigenvectors of the one-dimensional periodic problem on the fine grid. Constructing the unitary matrix

$$W = \left[\mathbf{w}_1^f\ \mathbf{w}_{1+N_c}^f\ \mathbf{w}_2^f\ \mathbf{w}_{2+N_c}^f\ \ldots\ \mathbf{w}_{n_c}^f\ \mathbf{w}_{n_c+N_c}^f\ \mathbf{w}_{N_c}^f\ \mathbf{w}_{N_f}^f\right],$$

it can be shown that

$$B_j^{per}W = W\hat{B}_j^{per}, \quad C_j^{per}W = W\hat{C}_j^{per}, \qquad j = 1, \ldots, n_f$$

where each of $\hat{B}_j^{per}$, $\hat{C}_j^{per}$ is an $N_f \times N_f$ block diagonal matrix with 2×2 blocks (see the appendix for details). Combining these according to (2.16) gives

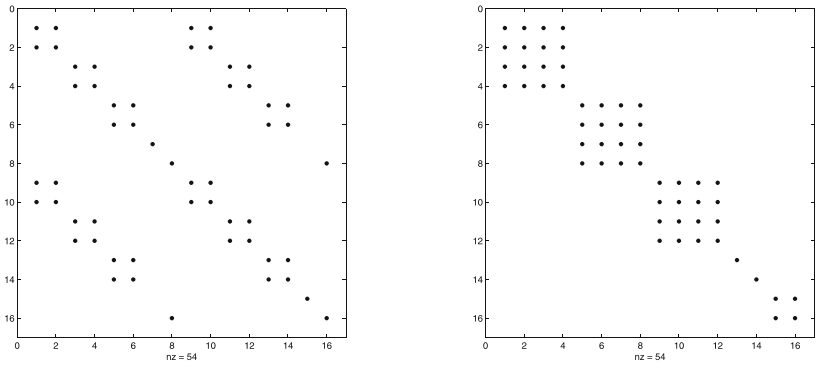
$$\mathcal{M}_j^{per} = \begin{bmatrix} \hat{B}_j^{per} & \hat{C}_j^{per} \\ \hat{C}_k^{per} & \hat{B}_k^{per} \end{bmatrix}, \quad j = 1, \ldots, n_c,\ k = N_f - j,$$

and a final permutation can be applied to produce matrices $\hat{\mathcal{M}}_j^{per}$ which are block-diagonal with 4×4 blocks. Representative sparsity patterns for $\mathcal{M}_j^{per}$ and $\hat{\mathcal{M}}_j^{per}$ are shown in Figure 3.1. As W is unitary,

$$\|\mathcal{M}_j^{per}\|_2 = \|\hat{\mathcal{M}}_j^{per}\|_2 = \max_{q=1,\ldots,N_c} \|X_{j,q}\|_2,$$

where $X_{j,q}$ is the 4×4 block of $\hat{\mathcal{M}}_j^{per}$ corresponding to the pair of eigenvectors $\mathbf{w}_q$, $\mathbf{w}_{q+N_c}$. From (2.18), it follows that

$$\|M^{per}\|_2 = \max\Big(\max_{j=1,\ldots,n_c}\Big[\max_{q=1,\ldots,N_c}\|X_{j,q}\|_2\Big], \|\mathcal{M}_{N_c}^{per}\|_2\Big), \tag{3.3}$$



(a) $\mathcal{M}_j^{per}$, $j = 1, \ldots, n_c$. (b) $\hat{\mathcal{M}}_j^{per}$, $j = 1, \ldots, n_c$.

Figure 3.1: Representative sparsity patterns of transformed subblocks $\mathcal{M}_j^{per}$ and $\hat{\mathcal{M}}_j^{per}$ of M^{per}, $N_f = 8$.

where

(3.4)

$$X_{j,q} = \begin{bmatrix} \hat{B}^{per}_{j,q} & \hat{C}^{per}_{j,q} \\ \hat{C}^{per}_{k,q} & \hat{B}^{per}_{k,q} \end{bmatrix} =$$

$$\begin{bmatrix} s^{\nu}_{1j}\left(1-\frac{f^2_{1j}d_{1j}}{g_j}\right) & -\frac{f_{1j}f_{2j}d_{2j}s^{\nu}_{2j}}{g_j} & -\frac{f_{1j}f_{1k}d_{1k}s^{\nu}_{1k}}{g_j} & -\frac{f_{1j}f_{2k}d_{2k}s^{\nu}_{2k}}{g_j} \\ -\frac{f_{2j}f_{1j}d_{1j}s^{\nu}_{1j}}{g_j} & s^{\nu}_{2j}\left(1-\frac{f^2_{2j}d_{2j}}{g_j}\right) & -\frac{f_{2j}f_{1k}d_{1k}s^{\nu}_{1k}}{g_j} & -\frac{f_{2j}f_{2k}d_{2k}s^{\nu}_{2k}}{g_j} \\ -\frac{f_{1k}f_{1j}d_{1j}s^{\nu}_{1j}}{g_j} & -\frac{f_{1k}f_{2j}d_{2j}s^{\nu}_{2j}}{g_j} & s^{\nu}_{1k}\left(1-\frac{f^2_{1k}d_{1k}}{g_j}\right) & -\frac{f_{1k}f_{2k}d_{2k}s^{\nu}_{2k}}{g_j} \\ -\frac{f_{2k}f_{1j}d_{1j}s^{\nu}_{1j}}{g_j} & -\frac{f_{2k}f_{2j}d_{2j}s^{\nu}_{2j}}{g_j} & -\frac{f_{2k}f_{1k}d_{1k}s^{\nu}_{1k}}{g_j} & s^{\nu}_{2k}\left(1-\frac{f^2_{2k}d_{2k}}{g_j}\right) \end{bmatrix}.$$

For definitions of s_{1j}, d_{1j}, etc. see the appendix.

4 Semiperiodic analysis.

As stated previously, the above analysis is applicable to any discretisation technique which gives rise to a stencil of the form (2.3). In this section we focus on the application of finite element discretisation (1.2) with

$$\delta = \max\left(0, \frac{1}{2} - \frac{\epsilon}{h}\right), \quad \delta^c = \max\left(0, \frac{1}{2} - \frac{\epsilon}{\bar{h}}\right),$$

to model problem (2.1). The coefficients in stencil (2.3) for this discretisation using bilinear finite elements are given by

$$\begin{aligned} m_1 &= \tfrac{4}{3}\left(\delta h + 2\epsilon\right), & m_2 &= \tfrac{1}{3}\left(\delta h - \epsilon\right), \\ m_3 &= -\tfrac{1}{3}\left[\left(2\delta - 1\right)h + \epsilon\right], & m_4 &= -\tfrac{1}{12}\left[\left(2\delta - 1\right)h + 4\epsilon\right], \\ m_5 &= -\tfrac{1}{3}\left[\left(2\delta + 1\right)h + \epsilon\right], & m_6 &= -\tfrac{1}{12}\left[\left(2\delta + 1\right)h + 4\epsilon\right]. \end{aligned}$$

These formulae can be used to generate the fine grid eigenvalues (2.5) (and, with appropriate values δ^c and $\bar{h}$, their coarse grid equivalents) which can then in turn be used in (6.1) to generate the entries in the matrices $X_{j,q}$ given by (3.4) (which are listed in the appendix). Here we use these values to evaluate (3.3) for this specific discretisation.

First, we note that we can take advantage of some symmetry in the problem. We have that

$$j \in I_j \equiv \{1, 2, \ldots, n_c\} \Rightarrow \theta_j \in \left(0, \tfrac{\pi}{2}\right), \quad q \in I_q \equiv \{1, 2, \ldots, N_c\} \Rightarrow \theta_q \in \left(0, \tfrac{\pi}{2}\right],$$

and that any pair of angles α and $\alpha^* = \pi/2 - \alpha$ with $\alpha \in [0, \frac{\pi}{4}]$ satisfy

$$\sin(\alpha^*) = \cos(\alpha), \;\; \cos(\alpha^*) = \sin(\alpha), \;\; \sin(2\alpha^*) = \sin(2\alpha), \;\; \sin(4\alpha^*) = -\sin(4\alpha).$$

Hence, for an index pair $q \in \{0, 1, \ldots, \frac{N_c}{2}\}$ and $q^* = N_c - q$, we have symmetry in the formulae in (6.3) in the sense that

$$\begin{array}{lll} |d_{1j}(q)| = |d_{2j}(q^*)|, & |d_{2j}(q)| = |d_{1j}(q^*)|, & |d_{1k}(q)| = |d_{2k}(q^*)|, \\ |d_{2k}(q)| = |d_{1k}(q^*)|, & |s_{1j}(q)| = |s_{2j}(q^*)|, & |s_{2j}(q)| = |s_{1j}(q^*)|, \\ |s_{1k}(q)| = |s_{2k}(q^*)|, & |s_{2k}(q)| = |s_{1k}(q^*)|, & |f_{1j}(q)| = |f_{2j}(q^*)|, \\ |f_{2j}(q)| = |f_{1j}(q^*)|, & |f_{1k}(q)| = |f_{2k}(q^*)|, & |f_{2k}(q)| = |f_{1k}(q^*)|, \\ |g_j(q)| = |g_j(q^*)|. & & \end{array}$$

This symmetry means that, without loss of generality, we need only consider indices $q \in \bar{I}_q \equiv \{\frac{N_c}{2}, \ldots, N_c\}$, or $\theta_q \in [\frac{\pi}{4}, \frac{\pi}{2}]$.

We will deal with two distinct regimes of mesh Peclet number separately.

4.1 Case 1: $P_h \geq 1$.

When $P_h \geq 1$ (so δ is given by (1.4)), it can be observed from experiment that

$$\max_{q \in \bar{I}_q} \|X_{j,q}\|_2 \tag{4.1}$$

always occurs when $q = N_c$, that is, when $\theta_q = \pi/2$, so we need only find an expression for $\|X_{j,N_c}\|_2$. Using the notation

$$\mathcal{C}_j = \cos(\theta_j), \quad \mathcal{C}_{2j} = \cos(2\theta_j), \quad \mathcal{S}_{j2} = \sin\left(\frac{\theta_j}{2}\right), \quad \mathcal{C}_{j2} = \cos\left(\frac{\theta_j}{2}\right), \tag{4.2}$$

from (6.3) we have

$$f_{1j} = 0, \quad f_{2j} = 2\mathcal{C}_{j2}^2, \quad f_{1k} = 0, \quad f_{2k} = -2\mathcal{S}_{j2}^2$$

so (3.4) can be permuted to

$$X_{j,N_c} = \begin{bmatrix} s_{1j}^\nu & 0 & 0 & 0 \\ 0 & s_{1k}^\nu & 0 & 0 \\ 0 & 0 & s_{2j}^\nu \left(1 - \frac{4\mathcal{C}_{j2}^4 d_{2j}}{g_j}\right) & \frac{4\mathcal{C}_{j2}^2 \mathcal{S}_{j2}^2 d_{2k} s_{2k}^\nu}{g_j} \\ 0 & 0 & \frac{4\mathcal{C}_{j2}^2 \mathcal{S}_{j2}^2 d_{2j} s_{2j}^\nu}{g_j} & s_{2k}^\nu \left(1 - \frac{4\mathcal{S}_{j2}^4 d_{2k}}{g_j}\right) \end{bmatrix}$$

and

$$\|X_{j,N_c}\|_2 = \max\left\{\left|s_{1j}^\nu\right|, \left|s_{1k}^\nu\right|, \|\hat{X}\|_2\right\} \tag{4.3}$$

where

$$\hat{X} = \begin{bmatrix} s_{2j}^\nu \left(1 - \frac{4\mathcal{C}_{j2}^4 d_{2j}}{g_j}\right) & \frac{4\mathcal{C}_{j2}^2 \mathcal{S}_{j2}^2 d_{2k} s_{2k}^\nu}{g_j} \\ \frac{4\mathcal{C}_{j2}^2 \mathcal{S}_{j2}^2 d_{2j} s_{2j}^\nu}{g_j} & s_{2k}^\nu \left(1 - \frac{4\mathcal{S}_{j2}^4 d_{2k}}{g_j}\right) \end{bmatrix}.$$

Furthermore,

$$d_{2j} = 4\epsilon\mathcal{S}_{j2}^2, \quad d_{2k} = 4\epsilon\mathcal{C}_{j2}^2, \quad g_j = 16\epsilon\mathcal{S}_{j2}^2\mathcal{C}_{j2}^2$$

so

$$\hat{X} = \begin{bmatrix} s_{2j}^\nu(1-\mathcal{C}_{j2}^2) & s_{2k}^\nu\mathcal{C}_{j2}^2 \\ s_{2j}^\nu\mathcal{S}_{j2}^2 & s_{2k}^\nu(1-\mathcal{S}_{j2}^2) \end{bmatrix} = \begin{bmatrix} s_{2j}^\nu\mathcal{S}_{j2}^2 & s_{2k}^\nu\mathcal{C}_{j2}^2 \\ s_{2j}^\nu\mathcal{S}_{j2}^2 & s_{2k}^\nu\mathcal{C}_{j2}^2 \end{bmatrix}$$

and

$$\|\hat{X}\|_2 = \sqrt{2\big(s_{2j}^{2\nu}\mathcal{S}_{j2}^4 + s_{2k}^{2\nu}\mathcal{C}_{j2}^4\big)}.$$

Also, it can be shown that

$$s_{2j} = s_{2k} = -\tfrac{1}{5}, \quad |s_{1j}| < \tfrac{1}{5}, \quad |s_{1k}| < \tfrac{1}{5} \tag{4.4}$$

so

$$\|\hat{X}\|_2 = \frac{\sqrt{2}\sqrt{\mathcal{S}_{j2}^4 + \mathcal{C}_{j2}^4}}{5^\nu} = \frac{\sqrt{3+\cos(2\theta_j)}}{\sqrt{2}(5^\nu)} > \frac{1}{5^\nu}.$$

Hence

$$\max_{q=1\ldots,N_c} \|X_{j,q}\|_2 = \|X_{j,N_c}\|_2 = \max\big\{|s_{1j}|^\nu, |s_{1k}|^\nu, \|\hat{X}\|_2\big\} = \frac{\sqrt{3+\cos(2\theta_j)}}{\sqrt{2}(5^\nu)}.$$

This is maximised over $j \in I_j$ when $j = 1$, that is,

$$\max_{j=1\ldots,n_c}\Big[\max_{q=1\ldots,N_c}\|X_{j,q}\|_2\Big] = \frac{\sqrt{3+\cos(2\pi h)}}{\sqrt{2}(5^\nu)}.$$

To evaluate (3.3), it remains only to find

$$\max\left(\frac{\sqrt{3+\cos(2\pi h)}}{\sqrt{2}(5^\nu)}, \big\|\mathcal{M}_{N_c}^{per}\big\|_2\right)$$

where, from (2.17),

$$\mathcal{M}_{N_c}^{per} = \big(S_{N_c}^{per}\big)^\nu.$$

Using the results in the appendix, we see that

$$S_j^{per}\mathcal{W}_q = \mathcal{W}_q \begin{bmatrix} s_{1j} & 0 \\ 0 & s_{2j} \end{bmatrix}$$

so

$$\big\|M_{N_c}^{per}\big\|_2 = \big\|\big(S_{N_c}^{per}\big)^\nu\big\|_2 = \max\big\{|s_{1N_c}|^\nu, |s_{2N_c}|^\nu\big\} = \frac{1}{5^\nu}.$$

We therefore have an analytic expression for the norm of the two-grid iteration matrix when $P_h \geq 1$:

$$\|M^{per}\|_2 = \frac{\sqrt{3+\cos(2\pi h)}}{\sqrt{2}(5^\nu)}. \tag{4.5}$$

Note that as $h \to 0$,

$$\|M^{per}\|_2 \to \frac{\sqrt{2}}{5^\nu}. \tag{4.6}$$

4.2 Case 2: $P_h < 1$.

When $P_h < 1$ (so $\delta = 0$ and no streamline diffusion is added), it can be observed from experiment that the maximum in (4.1) always occurs at an extreme value of $\bar{I}_q$, that is, when $q = N_c/2$ or when $q = N_c$ (corresponding to $\theta_q = \pi/4$ and $\theta_q = \pi/2$, respectively). In this case, the complexity of the formulae makes it difficult to find an exact expression for $\|X_{j,q}\|_2$. We also observe, however, that the maximum in terms of j required for (3.3) occurs when $j = 1$ (that is, when $\theta_j = \theta_1 = \pi h \to 0$ as $h \to 0$). We therefore derive an approximate expression for (4.1) by replacing θ_j by 0 in (6.3), which simplifies the algebra considerably.

When $\theta_q = \pi/2$, we must identify (4.3) as before. With the above assumption of very small h, we have

$$\|\hat{X}\|_2 = \sqrt{2\left(s_{2j}^{2\nu}\mathcal{S}_{j2}^4 + s_{2k}^{2\nu}\mathcal{C}_{j2}^4\right)} \approx \sqrt{2}|s_{2k}|^\nu.$$

As it can be shown that this is bigger than $|s_{1j}|^\nu$ and $|s_{1k}|^\nu$, we have the approximation (when $\theta_j = 0$ and $\theta_q = \pi/2$)

$$\|X_{j,N_c}\|_2 \approx \sqrt{2}\left(\frac{1+P_h}{11-P_h}\right)^\nu. \tag{4.7}$$

When $\theta_q = \pi/4$, replacing θ_j by 0 gives

$$f_{1j} = 1, \quad f_{2j} = 1, \quad f_{1k} = 0, \quad f_{2k} = 0$$

so (3.4) becomes

$$X_{j,N_c/2} = \begin{bmatrix} s_{1j}^\nu\left(1 - \frac{d_{1j}}{g_j}\right) & -\frac{d_{2j}s_{2j}^\nu}{g_j} & 0 & 0 \\ -\frac{d_{1j}s_{1j}^\nu}{g_j} & s_{2j}^\nu\left(1 - \frac{d_{2j}}{g_j}\right) & 0 & 0 \\ 0 & 0 & s_{1k}^\nu & 0 \\ 0 & 0 & 0 & s_{2k}^\nu \end{bmatrix}$$

and

$$\|X_{j,N_c/2}\|_2 = \max\left\{\left|s_{1k}^\nu\right|, \left|s_{2k}^\nu\right|, \|\bar{X}\|_2\right\}$$

where

$$\bar{X} = \begin{bmatrix} s_{1j}^\nu\left(1 - \frac{d_{1j}}{g_j}\right) & -\frac{d_{2j}s_{2j}^\nu}{g_j} \\ -\frac{d_{1j}s_{1j}^\nu}{g_j} & s_{2j}^\nu\left(1 - \frac{d_{2j}}{g_j}\right) \end{bmatrix}.$$

As

$$d_{2j} = \bar{d}_{1j}, \quad s_{1j} = \bar{s}_{2j}, \qquad g_j \in \mathbb{R},$$

we have

$$\begin{aligned}\|\bar{X}\|_2 &= \sqrt{(\text{max. eigenvalue of } \bar{X}^* \bar{X})} \\ &= \left| s_{1j}^{\nu} \left(1 - \frac{d_{1j}}{g_j} \right) \right| + \left| \frac{d_{2j} s_{2j}^{\nu}}{g_j} \right| \\ &= |s_{1j}|^{\nu} \left(\left| 1 - \frac{d_{1j}}{g_j} \right| + \left| \frac{d_{2j}}{g_j} \right| \right).\end{aligned}$$

But

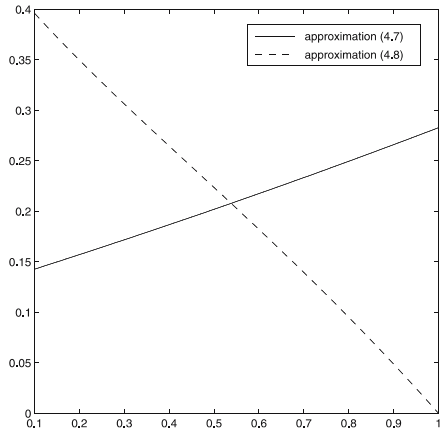
$$d_{1j} = 2\epsilon + hi, \quad d_{2j} = 2\epsilon - hi, \quad g_j = 4\epsilon, \quad |s_{ij}| = \frac{1 - P_h}{\sqrt{4 + (1 + P_h)^2}},$$

so

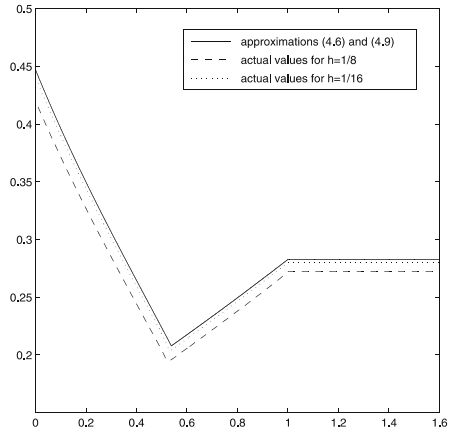
$$\|\bar{X}\|_2 = \left(\frac{1 - P_h}{\sqrt{4 + (1 + P_h)^2}} \right)^{\nu} \sqrt{1 + P_h^2}. \tag{4.8}$$

It can be shown that this is bigger than $|s_{1k}|^{\nu}$ and $|s_{2k}|^{\nu}$, so that (4.8) gives an approximation to $\|X_{j,N_c/2}\|_2$ (when $\theta_j = 0$ and $\theta_q = \pi/4$).

For a specific value of $P_h < 1$, the value of $\|M^{per}\|_2$ will be (approximately) equal to the maximum of (4.7) and (4.8). A plot of the two quantities against P_h is shown in Figure 4.1(a) ((4.7) solid line, (4.8) dashed line, $\nu = 1$). They are equal at a value $P_h^* \approx 0.538$, and



(a) Comparison of approximations (4.7) (solid) and (4.8) (dashed) for $P_h < 1$: the maximum of the two changes at $P_h = P_h^*$.



(b) Comparison of (4.6) and (4.9) (solid) with true values of $\|M^{per}\|_2$ for $h = 1/8$ (dashed) and $h = 1/16$ (dotted).

Figure 4.1: Assessment of the accuracy of approximations to $\|M^{per}\|_2$.

$$\|M^{per}\|_2 \approx \begin{cases} \left(\dfrac{1-P_h}{\sqrt{4+(1+P_h)^2}}\right)^{\nu} \sqrt{1+P_h^2} & P_h < P_h^*, \\ \sqrt{2}\left(\dfrac{1+P_h}{11-P_h}\right)^{\nu} & P_h^* < P_h < 1. \end{cases} \tag{4.9}$$

A plot of the full approximation to $\|M^{per}\|_2$, using (4.6) and (4.9) in the appropriate P_h regimes, is shown in Figure 4.1(b) (solid line). The actual values for $h = 1/8$ (dashed) and $h = 1/16$ (dotted) are also shown. For smaller values of h, the lines become indistinguishable from the approximation to the naked eye.

5 Comparisons with the Dirichlet problem.

As stated in the introduction, it is shown in [20] that applying the multigrid method we describe to the finite element discretisation under discussion results in iteration counts independent of the mesh size h. Similar results are obtained for the two-grid version studied here, as evidenced by the iteration counts in Table 5.1 (based on the stopping criterion $\|\mathbf{r}_k\|_2/\|\mathbf{r}_0\|_2 \leq 10^{-6}$). These results were obtained using Problem 3.1.1 from Elman et al. [7] which has Dirichlet boundary conditions $u(x,-1) = x$, $u(x,1) = 0$, $u(-1,y) = -1$ and $u(1,y) = 1$. The initial guess is the zero vector and the number of smoothing steps is $\nu = 1$ for this and all subsequent numerical examples (unless otherwise stated). The fact that the iteration count is decreasing as $P_h \to \infty$ is an artefact of our simple model problem: in this limit, the line Gauss-Seidel smoother becomes an exact solver for (2.1). It appears that the two-grid iteration works well for the Dirichlet problem (2.1) for all values of P_h: we would therefore expect to have a contraction number $\|M\|_2$ which is always less than less than one. It is clear from Figure 4.1(b) that this is the case for the semiperiodic problem analysed in the previous section. We now address the important question: are the results obtained from the semiperiodic analysis above useful for describing the actual behaviour for the full Dirichlet problem (2.1)?

Table 5.1: Two-grid iteration counts.

	ϵ										
h	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$	$\frac{1}{2048}$
$\frac{1}{4}$	5	5	5	5	5	4	4	3	2	2	2
$\frac{1}{8}$	7	7	6	6	5	5	4	4	3	2	2
$\frac{1}{16}$	7	7	7	6	5	5	5	4	4	3	2
$\frac{1}{32}$	7	7	7	7	6	5	5	4	4	3	3
$\frac{1}{64}$	7	7	7	7	6	5	5	4	4	4	3
$\frac{1}{128}$	7	6	6	6	6	6	5	4	4	4	3

To answer this, we first compare the theoretical bounds (4.6) and (4.9) for the semiperiodic problem with the value of $\|M\|_2$ for the Dirichlet problem, calculated numerically from the explicit formula (2.18), over a range of values of P_h. The results of this comparison for $0 < P_h \leq 1.6$ are illustrated in Figure 5.1(a). This shows the semiperiodic approximation (solid line) and the Dirichlet problem norms (dashed lines) for $h = 1/8, 1/16, \ldots, 1/512$, where the line for $h = 1/8$ is the lowermost and that for $h = 1/512$ the uppermost. There is an obvious distinction between the regimes $0 < P_h < 1$ and $P_h \geq 1$.

When $P_h \geq 1$, the Dirichlet behaviour is very well represented by the semiperiodic analysis, that is, the Dirichlet problem norm also tends to a constant as