University of Maryland
Institute for Advanced Computer Studies
Department of Computer Science

AN ANALYSIS OF SMOOTHING EFFECTS OF UPWINDING STRATEGIES FOR THE CONVECTION-DIFFUSION EQUATION

HOWARD C. ELMAN *AND ALISON RAMAGE †

CS-TR #4160 / UMIACS TR #2000-50
July 2000

Abstract

Using a technique for constructing analytic expressions for discrete solutions to the convection-diffusion equation, we examine and characterise the effects of upwinding strategies on solution quality. In particular, for grid-aligned flow and discretisation based on bilinear finite elements with streamline upwinding, we show precisely how the amount of upwinding included in the discrete operator affects solution oscillations and accuracy when boundary layers are present. In addition, we show that the same analytic techniques provide insight into other discretisations, such as a finite difference method that incorporates streamline diffusion, and the isotropic artificial diffusion method.

Keywords
Convection-diffusion equation, oscillations, Galerkin finite element method, streamline diffusion.

AMS subject classifications
65N22, 65N30, 65Q05, 35J25.

*Department of Computer Science, and Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742, USA. The work of this author was supported by National Science Foundation grant DMS9972490.

†Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, Scotland. The work of this author was supported by the Leverhulme Trust.

AN ANALYSIS OF SMOOTHING EFFECTS OF UPWINDING STRATEGIES FOR THE CONVECTION-DIFFUSION EQUATION

HOWARD C. ELMAN* AND ALISON RAMAGE†

Abstract

Using a technique for constructing analytic expressions for discrete solutions to the convection-diffusion equation, we examine and characterise the effects of upwinding strategies on solution quality. In particular, for grid-aligned flow and discretisation based on bilinear finite elements with streamline upwinding, we show precisely how the amount of upwinding included in the discrete operator affects solution oscillations and accuracy when boundary layers are present. In addition, we show that the same analytic techniques provide insight into other discretisations, such as a finite difference method that incorporates streamline diffusion, and the isotropic artificial diffusion method.

Keywords
Convection-diffusion equation, oscillations, Galerkin finite element method, streamline diffusion.

AMS subject classifications
65N22, 65N30, 65Q05, 35J25.

1. Introduction.

There are many discretisation strategies available for the linear convection-diffusion equation

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

where the small parameter ϵ and divergence-free convective velocity field $\mathbf{w} = (w_1(x,y), w_2(x,y))$ are given. In this paper, we analyse some well-known methods which involve the addition of upwinding to stabilise the discretisation for problems involving boundary layers. In particular, we focus on characterising exactly how this upwinding affects the resulting discrete solutions.

One possible discretisation technique is the Galerkin finite element method (see for example [3], [5], [6]). This is based on seeking a solution u of the weak form of equation (1.1),

$$\epsilon(\nabla u, \nabla v) + (\mathbf{w}.\nabla u, v) = (f, v) \qquad \forall\, v \in V,$$

where the test functions v are in the Sobolev space $V = \mathcal{H}_0^1(\Omega)$. Restricting this to a finite-dimensional subspace V_h of V gives

$$\epsilon(\nabla u_h, \nabla v) + (\mathbf{w}.\nabla u_h, v) = (f_h, v) \qquad \forall\, v \in V_h \tag{1.2}$$

*Department of Computer Science, and Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742, USA. The work of this author was supported by National Science Foundation grant DMS9972490.

†Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, Scotland. The work of this author was supported by the Leverhulme Trust.

where f_h is the $L^2(\Omega)$ orthogonal projection of f into V_h and h is a discretisation parameter. Choosing the test functions equal to a set of basis functions for V_h (usually continuous piecewise polynomials with local support) leads to a sparse linear system whose solution can be used to recover the discrete solution u_h.

One quantity which has an important effect on the quality of the resulting discrete solution is the mesh Péclet number

$$P_e = \frac{\|\mathbf{w}\| h}{2\epsilon}.$$

In particular, if $P_e > 1$, then the discrete solution obtained from the Galerkin method may exhibit non-physical oscillations. For the one-dimensional analogue of (2.1), this is well understood (see for example [6, p. 14]); for an analysis of the two-dimensional case, see [1]. An approach for minimising the deleterious effects of these oscillations, especially in areas of the domain away from boundary layers, is to stabilise the discrete problem by using an upwind discretisation. A particularly effective implementation of this idea is via the streamline diffusion method [5, §9.7], whereby a stabilisation parameter

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

is introduced (with $\delta > 0$ a parameter to be chosen) and the weak form (1.2) is replaced by

$$\epsilon(\nabla u_h, \nabla v) + (\mathbf{w}.\nabla u_h, v) + (\mathbf{w}\cdot\nabla u_h, \alpha\mathbf{w}\cdot\nabla v) = (f_h, v + \alpha\mathbf{w}\cdot\nabla v) \qquad \forall v \in V_h. \tag{1.4}$$

Formulation (1.4) has additional coercivity in the local flow direction, resulting in improved stability. Setting $\delta = 0$ reduces (1.4) to the standard Galerkin case (1.2).

In [1], we developed an analytic technique for characterising the nature of oscillations in discrete solutions arising from the Galerkin discretisation (1.2). More specifically, for the case of grid-aligned flow, we presented an analytic representation of the discrete solution, enabling isolation of any oscillatory behaviour in the direction of the flow. Using this framework, we studied the dependence of solution behaviour on the mesh Péclet number in some detail.

In this paper, we apply the tools developed in [1] to various upwinding strategies for discretising (1.1). For the most part, we focus on the streamline diffusion method (1.4), examining the effect of δ on the quality of the resulting discrete solutions. In section 2, we summarise the Fourier analysis presented in [1] and derive an explicit formula for the discrete streamline diffusion solution for a model problem with grid-aligned flow. Section 3 contains the details of this process in the case of bilinear finite elements. The resulting formulae allow us to investigate various issues which influence the choice of stabilisation parameter δ. We completely characterise the effect of δ on oscillations in the discrete solution in the flow direction, and discuss the implications of this for solution accuracy. In the remaining section we illustrate how the same approach can be used to understand other discretisation methods. We analyse an analogous streamline diffusion (upwind) discretisation for a finite difference stencil, and explain the comparative lack of effectiveness of isotropic artificial diffusion.

2. Summary of Fourier analysis. In this section, we summarise the Fourier techniques used in [1] to construct an analytic expression for the entries in the discrete solution vector $\mathbf{u}$.

Setting $\mathbf{w} = (0, 1)$ and f=0 in (1.1), we obtain the 'vertical wind' model problem

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

with Dirichlet boundary conditions as shown in Figure 2.1. Using a natural ordering of the un-

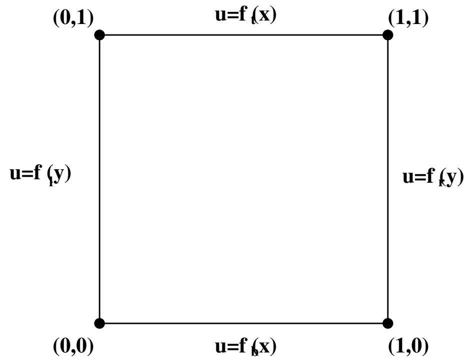


Figure 2.1: Boundary conditions.

knowns on a uniform grid of square bilinear elements with $N = 1/h$ elements in each dimension, both (1.2) and (1.4) give rise to a linear system

$$A\mathbf{u} = \mathbf{f} \tag{2.2}$$

where the coefficient matrix A is of order $(N-1)^2$. Denoting the coefficients of the computational molecule by

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

the matrix A can be written as

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

where $M_1 = \text{tridiag}(m_2, m_1, m_2)$, $M_2 = \text{tridiag}(m_4, m_3, m_4)$ and $M_3 = \text{tridiag}(m_6, m_5, m_6)$ are all tridiagonal matrices of order $N-1$. Given that the eigenvalues and eigenvectors of the blocks of A satisfy

$$\begin{array}{ll} M_1\mathbf{v}_j = \lambda_j\mathbf{v}_j & \lambda_j = m_1 + 2m_2\cos\frac{j\pi}{N} \\ M_2\mathbf{v}_j = \sigma_j\mathbf{v}_j & \sigma_j = m_3 + 2m_4\cos\frac{j\pi}{N} \\ M_3\mathbf{v}_j = \gamma_j\mathbf{v}_j & \gamma_j = m_5 + 2m_6\cos\frac{j\pi}{N} \end{array} \tag{2.5}$$

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

$$\mathbf{v}_j = \sqrt{\frac{2}{N}}\left[\sin\frac{j\pi}{N}, \quad \sin\frac{2j\pi}{N}, \quad \ldots, \sin\frac{(N-1)j\pi}{N}\right]^T, \tag{2.6}$$

we may obtain the decomposition

$$A = (\mathcal{V}P)T(\mathcal{V}P)^T$$

where $\mathcal{V} = \text{diag}(V, V, \ldots, V)$ is a block diagonal matrix with each block V having the $N-1$ eigenvectors (2.6) as its columns, and P is a permutation matrix of order $(N-1)^2$. The matrix

T is also block diagonal, with diagonal blocks $T_i = \text{tridiag}(\gamma_i, \lambda_i, \sigma_i)$, $i = 1, \ldots, N-1$. Using this decomposition, (2.2) implies

$$\mathbf{u} = \mathcal{V} P \mathbf{y} \tag{2.7}$$

where the vector $\mathbf{y}$ is the solution to the linear system

$$T\mathbf{y} = P^T \mathcal{V}^T \mathbf{f} \equiv \hat{\mathbf{f}}. \tag{2.8}$$

As T is block diagonal, this system can be partitioned into $N-1$ independent systems of the form

$$T_i \mathbf{y}_i = \hat{\mathbf{f}}_i \tag{2.9}$$

where T_i is defined above and $\mathbf{y}$ and $\hat{\mathbf{f}}$ are partitioned in the obvious way. Because T_i is a Toeplitz matrix, each of these systems can be considered as a three-term recurrence relation which can be solved analytically to give an expression for each entry y_{ik} of $\mathbf{y}_i$, $k = 1, \ldots, N-1$ in (2.9). Finally, to obtain an explicit formula for the entries of $\mathbf{u}$, we permute and transform these entries via (2.7) to get

$$u_{jk} = \sqrt{\frac{2}{N}} \sum_{i=1}^{N-1} \sin \frac{ij\pi}{N} y_{ik} \tag{2.10}$$

for $j, k = 1, \ldots, N-1$.

To obtain an expression for the entries y_{ik} in (2.10), we must consider the vectors $\hat{\mathbf{f}}_i$. As $f = 0$ in (2.1), the only nonzero entries in the original right-hand side vector $\mathbf{f}$ in (2.2) involve sums of certain matrix coefficients times boundary values, which are transformed and permuted to obtain $\hat{\mathbf{f}}$ in (2.8). The details of this process can be found in [1]: here we simply state that each right-hand side vector $\hat{\mathbf{f}}_i$, $i = 1, \ldots, N-1$, in (2.9) can be written as

$$\hat{\mathbf{f}}_i = \begin{bmatrix} \bar{b}_i + \bar{s}_i \\ \bar{s}_i \\ \vdots \\ \bar{s}_i \\ \bar{t}_i + \bar{s}_i \end{bmatrix}_{N-1}$$

where $\bar{b}_i$ involves data from the bottom boundary values, $\bar{t}_i$ involves data from the top boundary values and $\bar{s}_i$ combines information from the left and right boundary values. We will make the same assumption as in [1] that the functions $f_l(\mathbf{y})$ and $f_r(\mathbf{y})$ on the left and right boundaries are constant. This simplifies the presentation of the analysis.

The solution of each system (2.9) is now the solution of a three-term recurrence relation with constant coefficients whose auxiliary equation has roots

$$\mu_1(i) = \frac{-\lambda_i + \sqrt{\lambda_i^2 - 4\sigma_i\gamma_i}}{2\sigma_i}, \qquad \mu_2(i) = \frac{-\lambda_i - \sqrt{\lambda_i^2 - 4\sigma_i\gamma_i}}{2\sigma_i}. \tag{2.11}$$

The solution of this recurrence relation can be written as

$$y_{ik} = F_3(i) + [F_1(i) - F_3(i)]\, G_1(i,k) + [F_2(i) - F_3(i)]\, G_2(i,k) \tag{2.12}$$

where

$$\begin{aligned} G_1(i,k) &= \frac{\mu_1^k - \mu_2^k}{\mu_1^N - \mu_2^N}, \\ G_2(i,k) &= (1 - \mu_1^k) - (1 - \mu_1^N) \left[\frac{\mu_1^k - \mu_2^k}{\mu_1^N - \mu_2^N} \right], \end{aligned}$$

and the functions

$$F_1(i) = -\frac{\bar{t}_i}{\sigma_i}, \qquad F_2(i) = \frac{\bar{s}_i}{\sigma_i + \lambda_i + \gamma_i}, \qquad F_3(i) = -\frac{\bar{b}_i}{\gamma_i}$$

involve the coefficient matrix entries and boundary condition information (see [1] for details).

We emphasise that the functions $F_m(i)$, $m = 1, 2, 3$ in (2.12) are independent of the vertical grid index k: for fixed i, the behaviour of $\mathbf{y}$ in the streamline (vertical) direction depends only on the functions $G_1(i,k)$ and $G_2(i,k)$. In addition, as $F_1(i)$ is related to the top boundary values, $F_2(i)$ is related to the sum of the left and right boundary values (which have been assumed to be constant for this analysis) and $F_3(i)$ is related to the bottom boundary values, (2.12) shows that different boundary conditions will dictate how the functions $G_1(i,k)$ and $G_2(i,k)$ combine to produce different two-dimensional recurrence relation solutions y_{ik}. In the next section, we analyse the behaviour of these solutions in some detail for the streamline diffusion finite element discretisation (1.4) with bilinear elements.

3. Streamline diffusion discretisation with bilinear finite elements. In [1], an explicit expression for (2.12) for the Galerkin finite element method with bilinear elements was derived and analysed. Here we present the equivalent analysis for the streamline diffusion finite element discretisation (1.4) with a view to precisely characterising the effect of the extra diffusion on the oscillations that occur with the Galerkin method when $P_e > 1$. We again use bilinear elements.

3.1. The recurrence relation solution. The coefficients in stencil (2.3) for a streamline diffusion discretisation using bilinear finite elements are given by

$$m_1 = \tfrac{4}{3}\,(\delta h + 2\epsilon), \qquad m_2 = \tfrac{1}{3}\,(\delta h - \epsilon), \qquad m_3 = -\tfrac{1}{3}\,[(2\delta - 1)\,h + \epsilon],$$

$$m_4 = -\tfrac{1}{12}\,[(2\delta - 1)\,h + 4\epsilon], \quad m_5 = -\tfrac{1}{3}\,[(2\delta + 1)\,h + \epsilon], \quad m_6 = -\tfrac{1}{12}\,[(2\delta + 1)\,h + 4\epsilon].$$

For convenience, we introduce the notation

$$C_i = \cos\frac{i\pi}{N}$$

and write the eigenvalues (2.5) as

$$\begin{aligned}
\gamma_i &= \frac{1}{6}\{-2[\delta h(2 + C_i) + \epsilon(1 + 2C_i)] - h(2 + C_i)\} \\
\lambda_i &= \frac{2}{3}\{[\delta h(2 + C_i) + \epsilon(1 + 2C_i)] + 3\epsilon(1 - C_i)\} \\
\sigma_i &= \frac{1}{6}\{-2[\delta h(2 + C_i) + \epsilon(1 + 2C_i)] + h(2 + C_i)\},
\end{aligned}$$

$i = 1, \ldots, N-1$. Substituting these into (2.11) gives the expressions

$$\mu_{1,2} = \frac{-2\delta - \left[\dfrac{4 - C_i}{2 + C_i}\right]\dfrac{1}{P_e} \pm \sqrt{1 + \dfrac{12\delta(1 - C_i)}{(2 + C_i)}\dfrac{1}{P_e} + \dfrac{3(5 + C_i)(1 - C_i)}{(2 + C_i)^2}\dfrac{1}{P_e^2}}}{-2\delta + 1 - \left[\dfrac{1 + 2C_i}{2 + C_i}\right]\dfrac{1}{P_e}} \tag{3.1}$$

for the auxiliary equation roots in (2.12).

3.2. Oscillations in the recurrence relation solution. We know from [1, Thm 5.1] that if $P_e > 1$, then the recurrence relation solution $\mathbf{y}$ and the related discrete solution $\mathbf{u}$ to the pure Galerkin problem (1.2) usually exhibit oscillations. In this section we address the question of how the streamline diffusion parameter δ can be chosen to eliminate oscillations in the recurrence relation solution $\mathbf{y}$. The issue of how this affects the resulting $\mathbf{u}$ will be discussed in section 3.3.

Theorem 3.1 *If $P_e > 1$, then for any value of $i \in S_N \equiv \{1, \ldots, N-1\}$ there exists a parameter*

$$\delta_i^c = \frac{1}{2}\left(1 - \left[\frac{1+2C_i}{2+C_i}\right]\frac{1}{P_e}\right) \tag{3.2}$$

such that $\delta > \delta_i^c$ implies that $G_1(i,k)$ and $G_2(i,k)$ in (2.12) are non-oscillatory functions of k.

Proof. We have

$$G_1(i,k) = \frac{\mu_1^k - \mu_2^k}{\mu_1^N - \mu_2^N} = \left[\frac{\left(\dfrac{\mu_1}{\mu_2}\right)^k - 1}{\left(\dfrac{\mu_1}{\mu_2}\right)^N - 1}\right]\mu_2^{k-N} = \Theta(i,k)\,\mu_2^{k-N}.$$

As $|\mu_1/\mu_2| < 1$, $\Theta(i,k)$ is always positive. Hence if μ_2 is negative, $G_1(i,k)$ alternates in sign as k goes from 1 to $N-1$, that is, $G_1(i,k)$ is oscillatory for fixed $i \in S_N$. From (3.1), the numerator of μ_2 is always negative so, for δ_i^c given by (3.2), we have the conditions

$$\begin{cases} \delta > \delta_i^c \quad \Rightarrow \quad \mu_2 > 0,\ G_1(i,k) \text{ is non-oscillatory} \\[2ex] \delta < \delta_i^c \quad \Rightarrow \quad \mu_2 < 0,\ G_1(i,k) \text{ is oscillatory} \end{cases}.$$

In addition, it can be shown that $0 < \mu_1 < 1$ so that if $G_1(i,k)$ is non-oscillatory, then $G_2(i,k) = (1-\mu_1^k) - (1-\mu_1^N)G_1(i,k)$ must also be non-oscillatory. □

Sample plots of $G_1(i,k)$ for various values of $i \in S_N$ when $N = 16$ and $P_e = 3.125$ are given in Figure 3.1. Only the right half of the range of k has been plotted in each case to magnify the area of interest. Each subplot shows the behaviour for three distinct values of δ, namely $\delta = 0.2$ (solid line, o), $\delta = 0.4$ (dotted line, $\Diamond$) and $\delta = 0.6$ (dashed line, $\triangle$). Given the relevant critical values $\delta_1^c \simeq 0.34$, $\delta_{N/2}^c \simeq 0.42$ and $\delta_{N-1}^c \simeq 0.65$ for this problem, the dependence of oscillations on the value of δ is clear. For $\delta = 0.2$ (that is, $\delta < \delta_i^c\ \forall i \in S_N$), all functions $G_1(i,k)$ are oscillatory; for $\delta = 0.4$, $G_1(1,k)$ is non-oscillatory (as $\delta > \delta_1^c$) and $G_1(N/2,k)$ is only very mildly oscillatory; for $\delta = 0.6$, only $G_1(N-1,k)$ is oscillatory (as $\delta > \delta_i^c$ for $i = 1, N/2$). Analogous behaviour is seen in Figure 3.2 for $G_2(i,k)$ with the same parameter values, although the oscillations here occur about the function $1-\mu_1^k$ rather than zero.

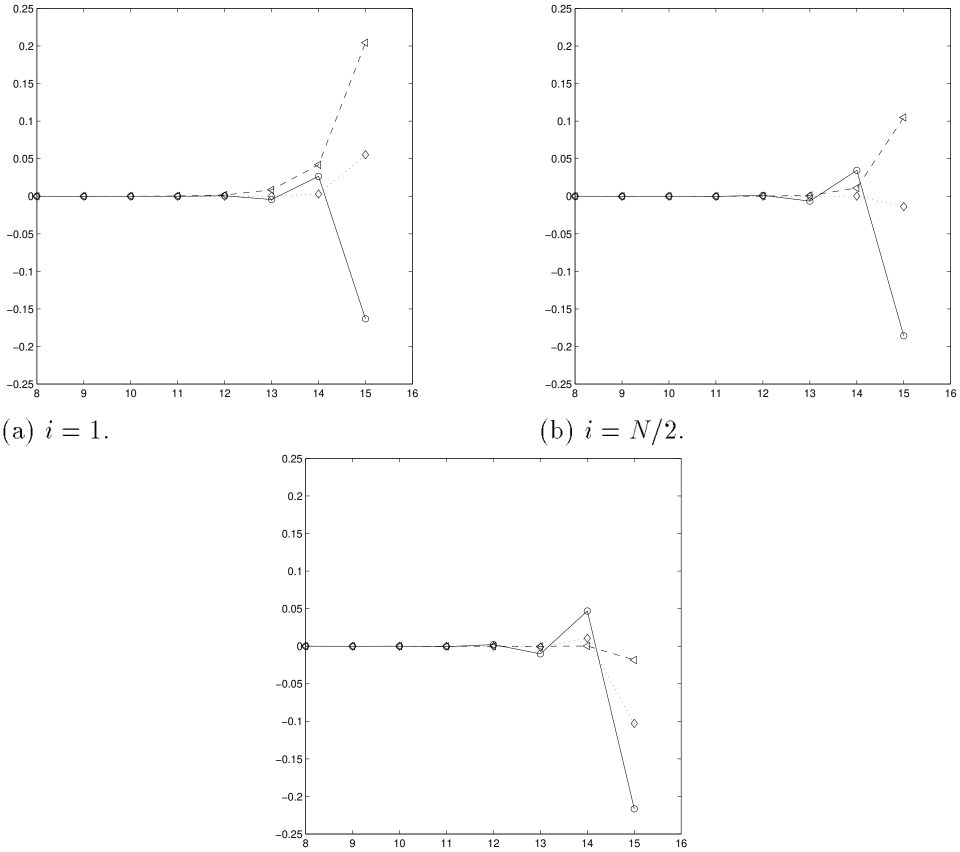
We now define

$$\delta_* = \frac{1}{2}\left(1 - \frac{1}{P_e}\right), \qquad \delta^* = \frac{1}{2}\left(1 + \frac{1}{P_e}\right) \tag{3.3}$$

(as in [2]), so that

$$\delta_* < \delta_i^c < \delta^* \tag{3.4}$$

for all values of $i \in S_N$. If $\delta \geq \delta^*$, then $\delta > \delta_i^c$ for each $i \in S_N$ and all of the functions $G_1(i,k)$ and $G_2(i,k)$ will be non-oscillatory in terms of k. We therefore have the following corollary to Theorem 3.1:



(a) $i = 1$.

(b) $i = N/2$.

(c) $i = N - 1$.

Figure 3.1: Plots of $G_1(i,k)$ against k for fixed i with $\delta = 0.2$ (solid, o), $\delta = 0.4$ (dotted, $\diamondsuit$) and $\delta = 0.6$ (dashed, $\triangle$).

Corollary 3.1 *For any value of δ such that $\delta \geq \delta^*$, the functions $G_1(i,k)$ and $G_2(i,k)$ in (2.12) are non-oscillatory functions of k for every $i \in S_N$. Hence the recurrence relation solution $\mathbf{y}$ is a sum of smooth functions and will not exhibit oscillations in the streamline direction.*

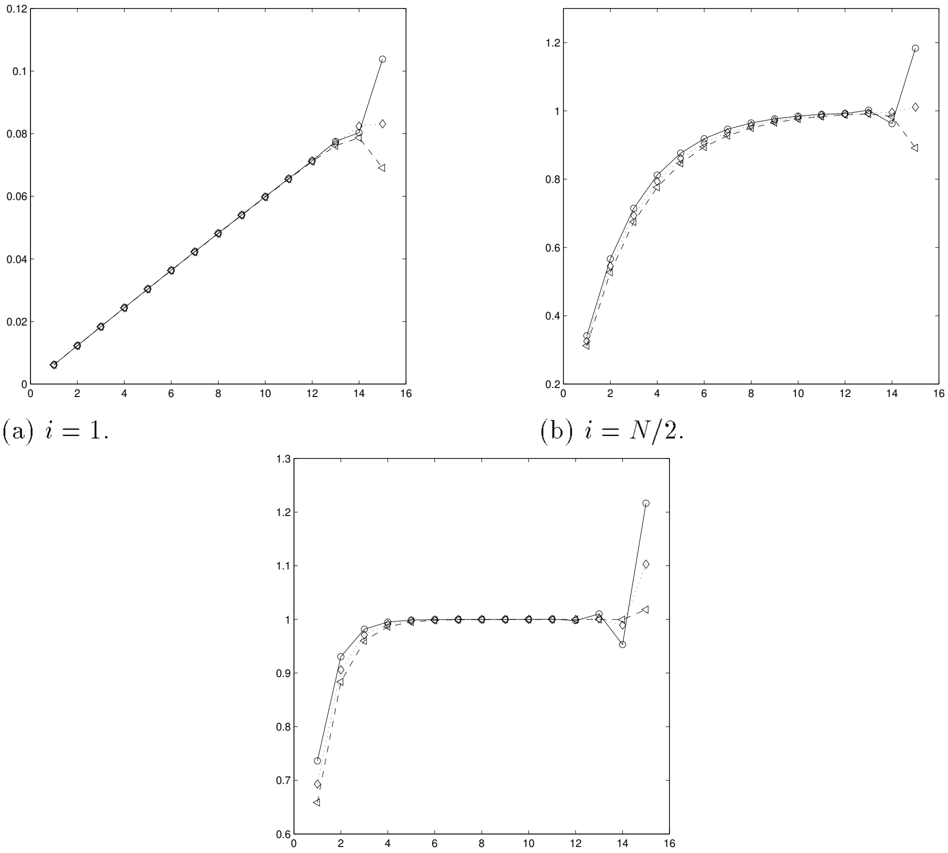
The case $\delta = \delta_i^c$ requires special attention. With this value, $\sigma_i = 0$ in (2.5) and the resulting matrix T_i in (2.9) is bidiagonal. This leads to a two-term recurrence relation with auxiliary equation root

$$\rho = \frac{1}{1 + \dfrac{3(1-C_i)}{2+C_i}\dfrac{1}{P_e}}$$

and solution

$$y_{ik} = F_3(i)\rho^k + F_2(i)(1-\rho^k). \tag{3.5}$$

As $0 < \rho < 1$ for any $i \in S_N$, y_{ik} is non-oscillatory in the streamline direction. In addition, $\rho \to 1$ as $P_e \to \infty$ giving the solution $y_{ik} = F_3(i)$. Looking ahead to section 3.3, applying transformation (2.10) gives $u_{jk} = f_b(x_j)$ (see (3.8)). This is the solution to the reduced problem (obtained by setting $\epsilon = 0$ in (2.1)) where the bottom boundary values are simply transported in the direction of the flow without any diffusion present. That is, with the choice $\delta = \delta_i^c$ for



(a) $i = 1$. (b) $i = N/2$. (c) $i = N - 1$.

Figure 3.2: Plots of $G_2(i,k)$ against k for fixed i with $\delta = 0.2$ (solid, o), $\delta = 0.4$ (dotted, $\diamondsuit$) and $\delta = 0.6$ (dashed, $\triangle$).

each i, the discrete solution is exact at every interior node in the limit as $P_e \to \infty$.

3.3. Oscillations in the discrete solution. In this section we consider the impact of transformation (2.10) on the recurrence relation solution $\mathbf{y}$, with a view to choosing δ to obtain an oscillation-free discrete solution $\mathbf{u}$. We begin by considering the functions $F_m(i)$, $m = 1, 2, 3$ in (2.12). Following the analysis of [1, §4.4 and Appendix] we can derive expressions

$$
\begin{aligned}
F_1(i) &= \sqrt{\frac{2}{N}} \sum_{p=1}^{N-1} f_t(x_p) \sin \frac{pi\pi}{N}, \\
F_2(i) &= f_l \sqrt{\frac{2}{N}} \sum_{p=1}^{N-1} \sin \frac{pi\pi}{N}, \\
F_3(i) &= \sqrt{\frac{2}{N}} \sum_{p=1}^{N-1} f_b(x_p) \sin \frac{pi\pi}{N},
\end{aligned}
\tag{3.6}
$$

for the streamline diffusion weight functions in the special case where the constant left and right boundary values f_l and f_r are equal. From (2.12), we therefore have

$$\begin{aligned} y_{ik} &= \sqrt{\frac{2}{N}}\sum_{p=1}^{N-1} f_b(x_p)\sin\frac{pi\pi}{N} + \sqrt{\frac{2}{N}}\sum_{p=1}^{N-1}[f_t(x_p)-f_b(x_p)]\sin\frac{pi\pi}{N}G_1(i,k) \\ &+ \sqrt{\frac{2}{N}}\sum_{p=1}^{N-1}[f_l-f_b(x_p)]\sin\frac{pi\pi}{N}G_2(i,k) \end{aligned} \tag{3.7}$$

[1, Thm 4.2]. Note that the expressions in (3.6) hold for any stencil of the form (2.3) whose entries sum to zero. In particular, this implies that the functions in (3.6) are the same for discretisations (1.2) and (1.4).

We now apply transformation (2.10) to (3.7) to obtain an expression for the entries of the discrete solution vector $\mathbf{u}$. As in [1], for the first term we have

$$\sqrt{\frac{2}{N}}\sum_{i=1}^{N-1}\sin\frac{ij\pi}{N}\left\{\sqrt{\frac{2}{N}}\sum_{p=1}^{N-1} f_b(x_p)\sin\frac{pi\pi}{N}\right\} = f_b(x_j) \tag{3.8}$$

where $f_b(\mathbf{x})$ is the bottom boundary function in Figure 2.1. Applying (2.10) to the full expression (3.7) therefore gives

$$u_{jk} = f_b(x_j) + \frac{2}{N}\sum_{i=1}^{N-1}[a_{ij}G_1(i,k) + b_{ij}G_2(i,k)] \tag{3.9}$$

where

$$\begin{aligned} a_{ij} &= \sin\frac{ij\pi}{N}\sum_{p=1}^{N-1}[f_t(x_p)-f_b(x_p)]\sin\frac{pi\pi}{N} \\ b_{ij} &= \sin\frac{ij\pi}{N}\sum_{p=1}^{N-1}[f_l-f_b(x_p)]\sin\frac{pi\pi}{N}. \end{aligned} \tag{3.10}$$

That is, along a streamline (j fixed), $\mathbf{u}$ consists of the bottom boundary value on that line plus a linear combination of the functions $G_1(i,k)$ and $G_2(i,k)$ for $i \in S_N$. Note that $a_{i(N-j)} = a_{ij}$ and $b_{i(N-j)} = b_{ij}$, so that if $f_b(\mathbf{x})$ is symmetric about the centre vertical line of the grid, then so is $\mathbf{u}$.

We can use the representation (3.9) to obtain insight into the effect of δ on the quality of the solution in the streamline direction. Recall from section 3.2 that if $\delta \geq \delta_i^c$ in (3.2) then the functions $G_1(i,k)$ and $G_2(i,k)$ are non-oscillatory in the streamline direction for that particular $i \in S_N$. In accordance with formulation (1.4), however, we would like to choose one global parameter δ for all values of $i \in S_N$. We can do this using Corollary 3.1, which implies that if $\delta \geq \delta^*$ in (3.3) then (3.9) is a sum of smooth functions, establishing the following result:

Theorem 3.2 *For a streamline diffusion discretisation of (2.1) with bilinear finite elements, the discrete solution* $\mathbf{u}$ *does not exhibit oscillations in the streamline direction when* $\delta \geq \delta^*$.

In practice, it turns out that the restriction on δ given by this theorem is too harsh, and it is possible to obtain a non-oscillatory $\mathbf{u}$ for values of δ smaller than δ^* due to the 'smoothing' nature of transformation (2.10). In [1], the precise effect of this transformation was studied in the context of the behaviour of the Galerkin finite element solution for different mesh Péclet numbers. Here we present a similar discussion of the effects of varying δ in the streamline diffusion method. We will use notation based on considering u_{jk} in (3.9) as a sum of smooth and oscillatory parts. That is, letting i^* be the lowest value of $i \in S_N$ such that $\delta < \delta_i^c$, we write

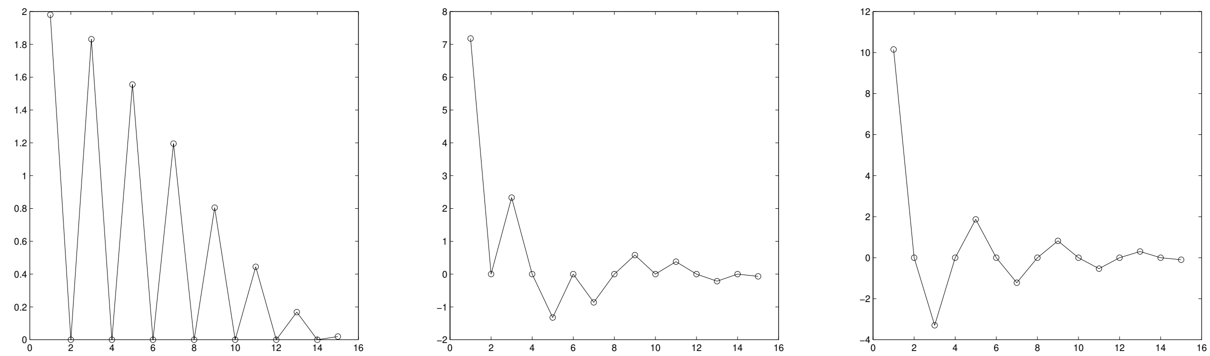
$$\begin{aligned} u_{jk} &= f_b(x_j) + \frac{2}{N}\left(\sum_{i=1}^{i^*-1}[a_{ij}G_1(i,k) + b_{ij}G_2(i,k)] + \sum_{i=i^*}^{N-1}[a_{ij}G_1(i,k) + b_{ij}G_2(i,k)]\right) \\ &= f_b(x_j) + S_{\text{smooth}} + S_{\text{osc}}. \qquad (3.11) \end{aligned}$$

Note that the above analysis implies $S_{\text{smooth}} = 0$ when $\delta \le \delta_*$ and $S_{\text{osc}} = 0$ when $\delta \ge \delta^*$. As δ increases from δ_*, i^* will increase so that S_{smooth} contains more and more of the terms, with the overall smoothness of $\mathbf{u}$ dependent on the relative size of the two sums S_{smooth} and S_{osc}.

We now focus on the specific example with boundary condition functions $f_t = 1$, $f_b = f_l = f_r = 0$. For this problem, the coefficients in (3.11) simplify to

$$a_{ij} = \sin\frac{ij\pi}{N}\sum_{p=1}^{N-1}\sin\frac{pi\pi}{N}, \qquad b_{ij} = 0,$$

with the magnitude of each a_{ij} decreasing rapidly as i goes from 1 to $N-1$ as shown in Figure 3.3 (taken from [1]). This means that the contributions to u_{jk} from the functions $G_1(i,k)$ are

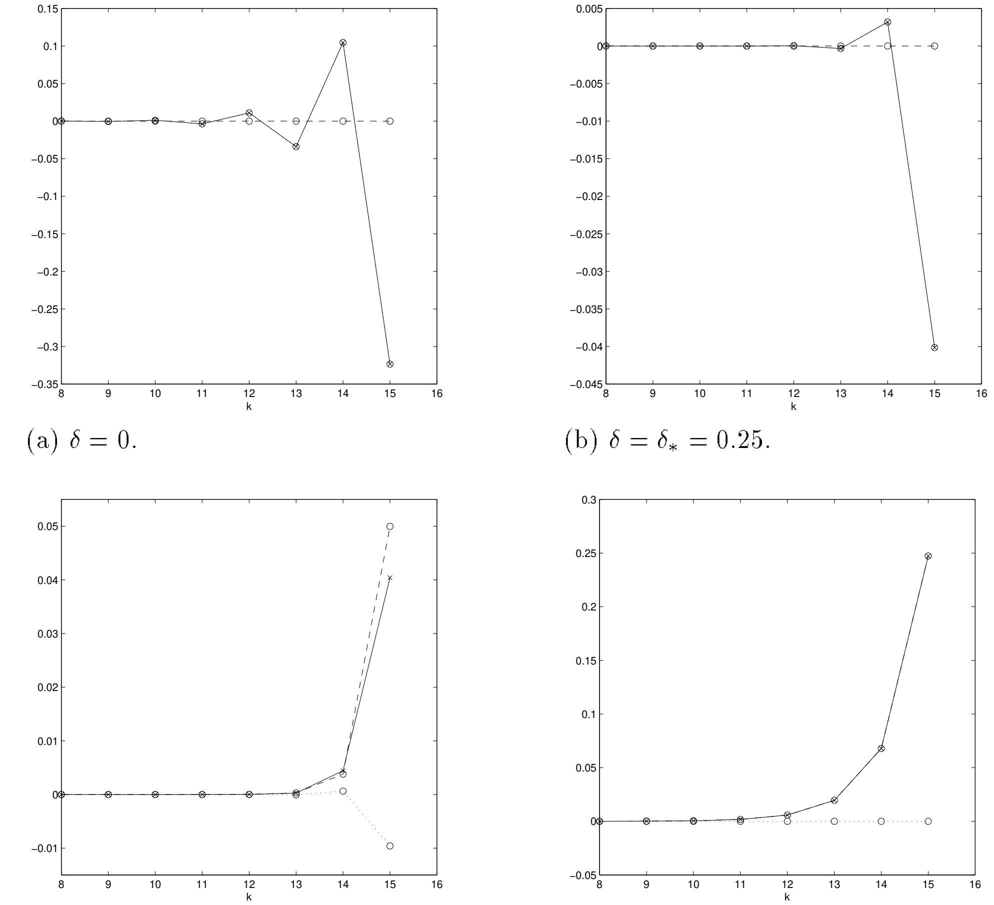


(a) $j = 1$. (b) $j = N/4$. (c) $j = N/2$.

Figure 3.3: Plots of coefficients a_{ij} against i for $N = 16$.

much larger for small indices i, so that the smoothness of $G_1(i,k)$ for small i plays a much more important role. In particular, it is not necessary for $G_1(i,k)$ to be non-oscillatory for all $i \in S_N$ in order for $|S_{\text{smooth}}|$ to dominate $|S_{\text{osc}}|$ and the resulting function $\mathbf{u}$ to be smooth.

We illustrate these ideas in Figures 3.4 and 3.5 for this example problem with $N = 16$ and $P_e = 2$. The first figure shows u_{1k} (or, equivalently, $u_{(N-1)k}$) plotted against k. This is the vertical cross-section of the solution obtained by fixing $j = 1$, which is the most oscillatory of the vertical cross-sections for this problem. Each plot shows a comparison of S_{smooth} (dotted line, o) and S_{osc} (dashed line, o) with u_{1k} (solid line, x) for a different value of δ, where again only the right half of the range of k has been plotted to magnify the area of interest. For this example, $\delta_* = 0.25$ and $\delta^* = 0.75$. Plot (a) shows the Galerkin case ($\delta = 0$) where all of the functions $G_1(i,k)$ are oscillatory and S_{smooth} is zero. This is still true in plot (b), where $\delta = \delta_*$, but the magnitude and extent of the oscillations has been reduced considerably. The result of choosing $\delta = \delta^*$ according to Theorem 3.2 to guarantee an oscillation-free discrete solution by ensuring a



(a) $\delta = 0$.

(b) $\delta = \delta_* = 0.25$.

(c) $\delta = \delta_{\mathrm{s}} = 0.354$.

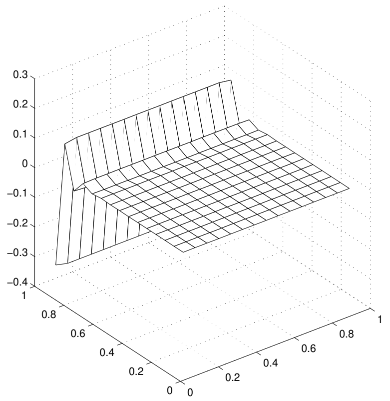
(d) $\delta = \delta^* = 0.75$.

Figure 3.4: Comparison of S_{smooth} (dotted line, o) and S_{osc} (dashed line, o) with u_{1k} (solid line, x).

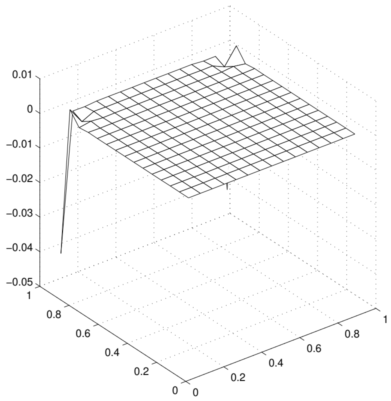
non-oscillatory $\mathbf{y}$ is shown in plot (d). Here too much extra diffusion has been added. Plot (c) shows u_{1k} for $\delta = \delta_{\mathrm{s}} = 0.354$, which lies in the interval (δ_7^c, δ_8^c), that is, $i^* = 8$. This is the lowest value of i^* such that S_{smooth} dominates (3.9) for this problem and u_{1k} is non-oscillatory.

The corresponding full two-dimensional solutions $\mathbf{u}$ are shown in Figure 3.5, where the boundary values have been omitted so that the fine detail of each solution is visible. The overall behaviour corresponds to that seen from the cross-sections: the severe oscillations present when $\delta = 0$ are almost eliminated by choosing $\delta = \delta_*$, and setting $\delta = \delta^*$ gives a smooth but overly-diffuse solution. For $\delta = \delta_{\mathrm{s}}$, the oscillations along the lines u_{1k} and $u_{(15)k}$ have just been eliminated to give a completely smooth solution in the flow direction.

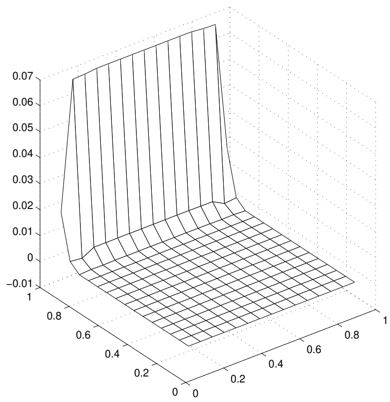
3.4. Solution accuracy. We have now completely characterised the effect of δ on oscillations in the flow direction. One important question which remains is how the choice of δ affects the accuracy of the discrete solution. To investigate this, we again focus on the example problem of the previous section with $f_t = 1$, $f_b = f_l = f_r = 0$. We compare solutions on a 16×16 grid with $\epsilon = 0.015625$ (so $P_e = 2$) with a reference solution for the same value of ϵ on a 256×256 grid. On this fine grid, we use the Galerkin method ($\delta = 0$) as $P_e = 0.125 \ll 1$ and there are no oscillations. In what follows, we will denote the fine grid nodal solution vector by



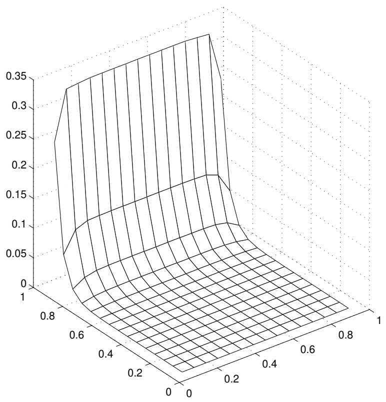
(a) $\delta = 0$.



(b) $\delta = \delta_* = 0.25$.



(c) $\delta = \delta_{\mathrm{s}} = 0.354$.



(d) $\delta = \delta^* = 0.75$.

Figure 3.5: Discrete solution at interior node points for $N = 16$, $P_e = 2$.

$\mathbf{u}_{256}$ and its associated finite element solution by u_{256}, likewise for the coarse grid solutions $\mathbf{u}_{16}^{\delta}$ and u_{16}^{δ}.

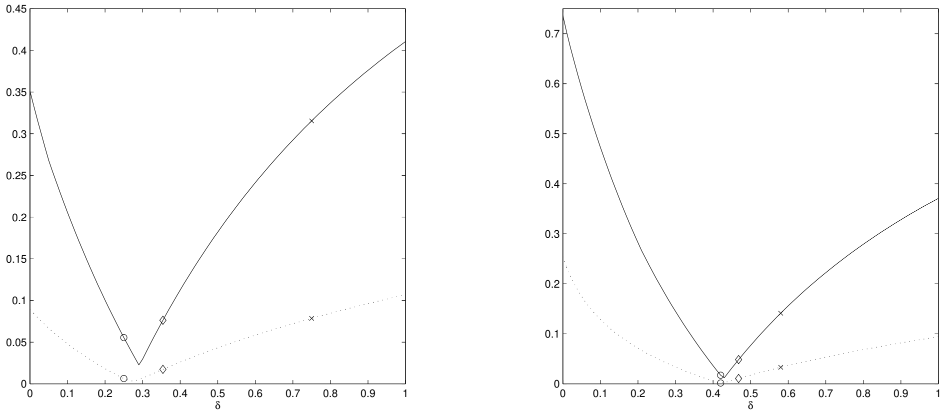
Figure 3.6(a) shows the variation with δ of the error measured in two different norms. In both cases the norm of the error is plotted against δ for $0 \leq \delta \leq 1$ with the values of δ_* (o), δ_{s} ($\diamondsuit$) and δ^* (x) highlighted. For $P_e = 6.25$, $\delta_* = 0.42$, $\delta_{\mathrm{s}} = 0.468$ and $\delta^* = 0.58$. The solid line represents the discrete $L_\infty[0,1]$ norm defined by

$$\|\mathbf{u}_{256} - \mathbf{u}_{16}^{\delta}\|_\infty = \max_{i,j} |\mathbf{u}_{256}(x_i, y_j) - \mathbf{u}_{16}^{\delta}(x_i, y_j)|$$

where the points $(x_i, y_j) = (ih, jh)$ are the nodes of the 16×16 grid. When using the finite element method, it may be more natural to work with the L_2 norm

$$\|u_{256} - u_{16}^{\delta}\|_2 = \left\{ \int_\Omega \left(u_{256} - u_{16}^{\delta} \right)^2 \right\}^{\frac{1}{2}}. \tag{3.12}$$

However, this measure leads to misleading results for singular perturbation problems of this type as the overall error is heavily dominated by the error in the boundary layer, which we cannot hope to resolve on a 16×16 uniform grid using low order elements. A more meaningful measure of the error for our purposes is obtained using the L_2 norm of the error away from the boundary



(a) $P_e = 2$.

(b) $P_e = 6.25$.

Figure 3.6: Error variation with δ in the discrete L_∞ norm (solid) and L_2 norm (dotted) for $N = 16$.

layer; that is, we omit the top row of coarse grid elements from the region of integration in (3.12) and integrate over $(0,1) \times (0, 0.9375)$ instead of $\Omega = (0,1) \times (0,1)$. This is the norm represented by the dotted line in Figure 3.6(a). We note in passing that this curve is very similar to that obtained for the discrete L_2 norm defined by

$$\|\mathbf{u}_{256} - \mathbf{u}_{16}^{\delta}\|_2 = \left\{ \sum_{i,j=0}^{N} \left(u_{256}(x_i, y_j) - u_{16}^{\delta}(x_i, y_j) \right)^2 \right\}^{\frac{1}{2}}$$

where (x_i, y_j) is again a node of the coarse grid.

We conclude this section with some observations about choosing δ in practice. From Figure 3.6, it is clear that the optimal choice of δ in terms of solution accuracy depends on the norm in which the error is measured. Setting $\delta = \delta_{\mathrm{s}}$, which produces a completely oscillation-free discrete solution $\mathbf{u}$, does not result in the most accurate solution. Also, δ_{s} is not in general readily determined. The minimum error occurs for δ close to δ_* in all cases and δ_* is defined by (3.3) for any P_e. In addition, δ_* coincides with the value shown in [2] to be good for efficient solution of the resulting linear system by the GMRES iterative method.

4. Applications to other discretisations. Analysis of this type can be applied to any discretisation whose stencil is of the form (2.3). We comment on two particular cases of interest here.

4.1. Finite differences with streamline diffusion. The usual central finite difference discretisation of (1.1) can also be stabilised using streamline diffusion, see for example [7, p. 1465]. Specifically, we apply the finite difference method to the differential equation

$$-(\epsilon \nabla^2 + \nabla \cdot D \nabla) u(x,y) + \mathbf{w} \cdot \nabla u(x,y) = f(x,y)$$

where diffusion in the streamline direction is added using

$$D = \alpha \begin{bmatrix} c^2 & cs \\ cs & s^s \end{bmatrix}$$

with

$$c = \frac{w_1}{\|\mathbf{w}\|}, \qquad s = \frac{w_2}{\|\mathbf{w}\|}$$

and α as in (1.3). Assuming for convenience that $\|\mathbf{w}\| = 1$, the full computational molecule is given by

$$\begin{array}{ccccc}
\dfrac{w_1 w_2 \delta}{2h} & & -\dfrac{\epsilon}{h^2} + \dfrac{w_2}{2h} - \dfrac{w_2^2\delta}{h} & & -\dfrac{w_1 w_2 \delta}{2h} \\
& \nwarrow & \uparrow & \nearrow & \\
-\dfrac{\epsilon}{h^2} - \dfrac{w_1}{2h} - \dfrac{w_1^2\delta}{h} & \leftarrow & \dfrac{4\epsilon}{h^2} + \dfrac{2\delta}{h} & \rightarrow & -\dfrac{\epsilon}{h^2} + \dfrac{w_1}{2h} - \dfrac{w_1^2\delta}{h} \\
& \swarrow & \downarrow & \searrow & \\
-\dfrac{w_1 w_2 \delta}{2h} & & -\dfrac{\epsilon}{h^2} - \dfrac{w_2}{2h} - \dfrac{w_2^2\delta}{h} & & \dfrac{w_1 w_2 \delta}{2h}
\end{array}.$$

This simplifies to a stencil of standard five-point type for our model problem (2.1) with grid-aligned flow. Using the notation of (2.3), the stencil entries are

$$m_1 = \frac{4\epsilon}{h^2} + \frac{2\delta}{h}, \qquad m_2 = -\frac{\epsilon}{h^2}, \qquad m_3 = -\frac{\epsilon}{h^2} + \frac{1}{2h} - \frac{\delta}{h},$$

$$m_4 = 0, \qquad m_5 = -\frac{\epsilon}{h^2} - \frac{1}{2h} - \frac{\delta}{h}, \qquad m_6 = 0$$

with related eigenvalues

$$\gamma_i = \frac{1}{h^2}\left[-(\epsilon + \delta h) - \frac{h}{2}\right], \qquad \lambda_i = \frac{1}{h^2}\left[2(\epsilon + \delta h) + 2\epsilon(1 - C_i)\right], \qquad \sigma_i = \frac{1}{h^2}\left[-(\epsilon + \delta h) + \frac{h}{2}\right].$$

This results in the expressions

$$\mu_{1,2} = \frac{-2\delta - [2 - C_i]\dfrac{1}{P_e} \pm \sqrt{1 + 4\delta(1 - C_i)\dfrac{1}{P_e} + (1 - C_i)(3 - C_i)\dfrac{1}{P_e^2}}}{-2\delta + 1 - \dfrac{1}{P_e}}.$$

for the roots of the recurrence relation which appear in (2.12).

Here the sign of μ_2 (and hence the nature of the corresponding functions $G_1(i,k)$ and $G_2(i,k)$, $i \in S_N$) is independent of i: as the numerator of μ_2 is always negative, we simply have the conditions

$$\begin{cases} \delta > \delta_* \;\Rightarrow\; \mu_2 > 0, \; G_1(i,k) \text{ is non-oscillatory} \\[2ex] \delta < \delta_* \;\Rightarrow\; \mu_2 < 0, \; G_1(i,k) \text{ is oscillatory} \end{cases},$$

where δ_* is given by (3.3). Hence the result equivalent to Theorem 3.1 is given by the following theorem:

Theorem 4.1 *For a streamline diffusion finite difference discretisation with $P_e > 1$, $\delta > \delta_*$ implies that $G_1(i,k)$ and $G_2(i,k)$ in (2.12) are non-oscillatory functions of k for any value of $i \in S_N$.*

The special case $\delta = \delta_*$ leads to the two-term recurrence with auxiliary equation root

$$\rho = \frac{1}{1 + \dfrac{(1 - C_i)}{P_e}}$$

and solution (3.5). Because $\rho > 1$, this solution is non-oscillatory in the streamline direction for all $i \in S_N$ and, as in the finite element case, tends to the nodally exact solution in the limit as $P_e \to \infty$.

The fact that there is one critical parameter (independent of i) here means that there is no issue about selecting a global parameter δ as we had in the finite element case. Furthermore, the analysis of the effect of transforming from $\mathbf{y}$ to $\mathbf{u}$ (cf. section 3.3) is greatly simplified. In particular, for the same specific example problem with $f_t = 1$ and $f_b = f_l = f_r = 0$ studied in section 3.3, the equivalent expression to (3.11) using finite differences has $S_{\text{smooth}} = 0$ when $\delta < \delta_*$ and $S_{\text{osc}} = 0$ when $\delta > \delta_*$. Thus we immediately have the following theorem (cf. Theorem 3.2):

Theorem 4.2 *For a streamline diffusion finite difference discretisation of (2.1), the discrete solution* $\mathbf{u}$ *does not exhibit oscillations in the streamline direction when* $\delta \geq \delta_*$.

That is, in contrast to the finite element case, there is no 'smoothing' introduced by the Fourier transformation: the same single parameter governs the presence of oscillations in both the recurrence relation solution $\mathbf{y}$ and the discrete two-dimensional solution $\mathbf{u}$.

4.2. Artificial diffusion. So far we have focused on adding smoothing in the streamline direction only, which is just one of the many stabilisation methods available. In this section we analyse the artificial diffusion method (see for example [4, pp. 218-219]) with a view to comparing its smoothing effect to that of streamline diffusion. The artificial diffusion method works by adding diffusion in an isotropic way which does not take account of flow direction, and it is well known that this can result in smearing of internal layers. We can use the analytical techniques presented in this paper to confirm that the streamline diffusion method avoids this problem.

We again consider a vertical wind model problem using bilinear finite elements on a uniform grid. The idea of the artificial diffusion method is to replace equation (2.1) by

$$-(\epsilon + \delta h)\nabla^2 u + \frac{\partial u}{\partial y} = 0 \qquad \text{in } \Omega = (0,1) \times (0,1), \tag{4.1}$$

with δ once again a stabilisation parameter to be chosen. Galerkin discretisation using bilinear finite elements results in a matrix is of the form (2.4), which is therefore covered by our analysis. The stencil entries in this case are given by

$$m_1 = \frac{8}{3}(\delta h + \epsilon), \qquad m_2 = -\frac{1}{3}(\delta h + \epsilon), \qquad m_3 = -\frac{1}{3}[(\delta - 1)h + \epsilon],$$

$$m_4 = -\frac{1}{12}[(4\delta - 1)h + 4\epsilon], \quad m_5 = -\frac{1}{3}[(\delta + 1)h + \epsilon], \quad m_6 = -\frac{1}{12}[(4\delta + 1)h + 4\epsilon],$$

so the roots (2.11) of the corresponding recurrence relation are given by

$$\mu_{1,2} = \frac{-\left(2\delta + \dfrac{1}{P_e}\right)\left[\dfrac{4 - C_i}{2 + C_i}\right] \pm \sqrt{1 + \dfrac{3(1 - C_i)(5 + C_i)}{(2 + C_i)^2}\left(2\delta + \dfrac{1}{P_e}\right)^2}}{1 - \left(2\delta + \dfrac{1}{P_e}\right)\left[\dfrac{1 + 2C_i}{2 + C_i}\right]}. \tag{4.2}$$

First we briefly consider the issue of oscillations in the streamline direction. Here, as in section 3.2, the sign of μ_2 (and hence the presence of oscillations in the recurrence relation solution) depends on the value of $i \in S_N$. Defining the new critical value

$$\tilde{\delta}_i^c = \frac{1}{2}\left(\left[\frac{2+C_i}{1+2C_i}\right] - \frac{1}{P_e}\right),$$

we have different conditions for two sets of i values, namely

$$1 \le i \le \tfrac{2}{3}N: \quad \begin{cases} \delta > \tilde{\delta}_i^c & \Rightarrow \quad \mu_2 > 0,\ G_1(i,k) \text{ is non-oscillatory} \\ \\ \delta < \tilde{\delta}_i^c & \Rightarrow \quad \mu_2 < 0,\ G_1(i,k) \text{ is oscillatory} \end{cases}$$

$$\tfrac{2}{3}N < i \le N-1: \quad \mu_2 < 0,\ G_1(i,k) \text{ is oscillatory.}$$

Notice that this is different from the streamline diffusion case (cf. Theorem 3.1) in that there is no choice of δ which will make the recurrence relation solution oscillation free, as some of the contributing functions $G_1(i,k)$ are always oscillatory. However, it can be seen using an argument of the type presented in section 3.3 that the transformed solution is again dominated by contributions from functions pertaining to lower values of i. Hence, despite the fact that $G_1(i,k)$ is always oscillatory for large i, it is still possible to obtain a non-oscillatory discrete solution $\mathbf{u}$. Note that inequality (3.4) is satisfied with δ_c^i replaced by $\tilde{\delta}_c^i$. For the particular (i-independent) choice $\delta = \delta_*$ from (3.3), equation (4.1) (and hence the artificial diffusion solution) is independent of ϵ.

To gain insight into the main difference between this method and the streamline diffusion technique, we must examine solution behaviour in the 'crosswind' direction, that is, perpendicular to the direction of the flow. To fix ideas, we will use the discontinuous boundary conditions

$$f_b(\mathbf{x}) = \begin{cases} 0, & x_i < 0.5 \\ 1, & x_i \ge 0.5 \end{cases}, \quad f_r(\mathbf{y}) = 1, \quad f_t(\mathbf{x}) = f_l(\mathbf{y}) = 0$$

so that the solution has an internal layer along $x = 0.5$ as well as a boundary layer along the right half of the top boundary. The internal layer derives from propagation of the bottom boundary condition through the domain and, as $\epsilon \to 0$, the width of this layer tends to zero. Ideally, this phenomenon should be reproduced by a discretisation method, that is, we would like to obtain a set of discrete solutions $\mathbf{u}$ in this limit whose variation from the bottom boundary function is independent of j for fixed k. We now show that while the streamline diffusion method has this property, the artificial diffusion method does not.

Consider the recurrence relation solution vector $\mathbf{y}$ for this problem. From (2.12), its entries are given by

$$y_{ik} = F_3(i)\,(1 - G_1(i,k)) + [F_2(i) - F_3(i)]\,G_2(i,k) \tag{4.3}$$

with

$$F_2(i) = \sqrt{\frac{2}{N}}\left[\frac{(-1)^{i+1}\sin\dfrac{i\pi}{N}}{2\left(1-\cos\dfrac{i\pi}{N}\right)}\right]$$

[1, Appendix] and $F_3(i)$ as in (3.6). As the functions $F_2(i)$ and $F_3(i)$ are the same for both discretisations, any difference in solution behaviour must come from a difference in the behaviour of the functions $G_1(i,k)$ and $G_2(i,k)$ associated with the two methods. We therefore now focus on how these functions vary with $i \in S_N$ as $\epsilon \to 0$ ($P_e \to \infty$) for $k \in S_N$ fixed. To simplify the presentation of this analysis, we will assume that δ is fixed independent of P_e, with $\delta \ne 0, 0.5$.

With the streamline diffusion discretisation, neglecting terms of $O(P_e^{-1})$ and higher in (3.1) gives the approximations

$$\mu_1 \simeq 1, \qquad \mu_2 \simeq \frac{2\delta+1}{2\delta-1} \equiv \beta$$

so that

$$\begin{aligned} G_1(i,k) &= \frac{\mu_1^k - \mu_2^k}{\mu_1^N - \mu_2^N} \simeq \frac{1-\beta^k}{1-\beta^N} \equiv G_1^a(k), \\ G_2(i,k) &= (1-\mu_1^k) - (1-\mu_1^N) G_1(i,k) \simeq 0. \end{aligned}$$

Thus, in the limit as $P_e \to \infty$, both functions are independent of i. We then have

$$y_{ik} \simeq F_3(i)(1 - G_1^a(k))$$

hence, using (2.7),

$$u_{jk} \simeq f_b(x_j)(1 - G_1^a(k)).$$

That is, the variation of u_{jk} from the bottom boundary function is independent of j in this limit. For the artificial diffusion discretisation, however, neglecting terms of $O(P_e^{-1})$ and higher in (4.2) gives

$$\mu_{1,2} \simeq \frac{-2\delta(4-C_i) \pm \sqrt{4(1+15\delta^2) + 4(1-12\delta^2)C_i + (1-12\delta^2)C_i^2}}{2(1-\delta) + (1-4\delta)C_i}$$

leading to approximations for $G_1(i,k)$ and $G_2(i,k)$ which depend on i through C_i. From (4.3) the solution is therefore

$$u_{jk} \simeq f_b(x_j) - \sqrt{\frac{2}{N}} \sum_{i=1}^{N-1} \sin\frac{ij\pi}{N} \left(F_3(i) G_1(i,k) - [F_2(i) - F_3(i)]\, G_2(i,k)\right).$$

This has a j-dependence which the continuous solution in this limit does not.

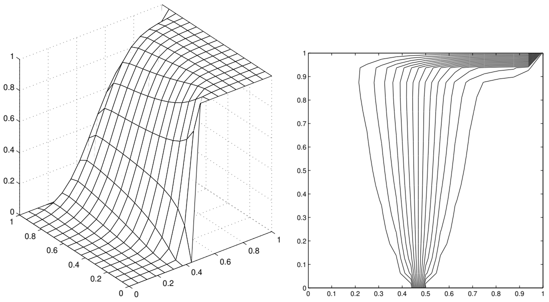
This fundamental difference between the discretisations is demonstrated pictorially in Figure 4.1, which shows streamline and artificial diffusion approximations (and associated contour plots) for this example problem with two values of ϵ, $\delta = 0.4$ and $N = 16$. Plots (a) and (b) show that the streamline diffusion method captures the narrowing of the internal layer exhibited by the continuous solution as $\epsilon \to 0$ ($P_e \to \infty$). The equivalent artificial diffusion approximation does not, as shown in plots (c) and (d).

References

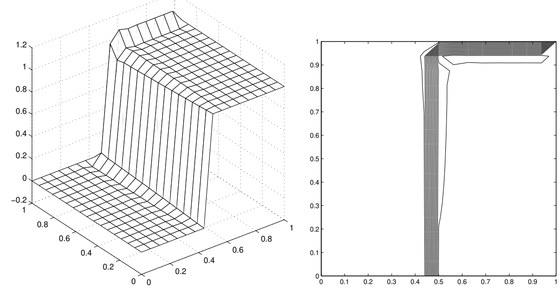
[1] H.C. Elman and A. Ramage. A characterisation of oscillations in the discrete two-dimensional convection-diffusion equation. Technical Report CS-TR-4118, University of Maryland, College Park MD 20742, 2000.

[2] B. Fischer, A. Ramage, D. J. Silvester, and A. J. Wathen. On parameter choice and iterative convergence for stabilised discretisations of advection-diffusion problems. *Comput. Methods Appl. Mech. Engrg*, 179:179–195, 1999.

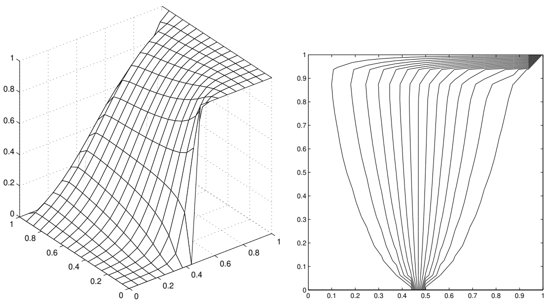
[3] P.M. Gresho and R.L. Sani. *Incompressible Flow and the Finite Element Method.* John Wiley and Sons, Chichester, 1999.



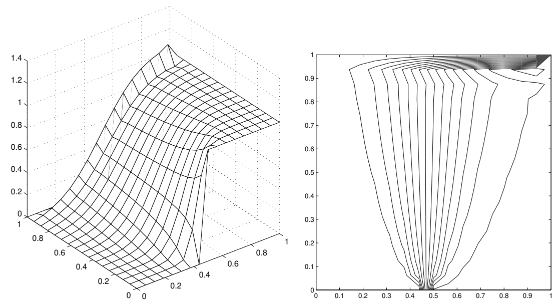
(a) Streamline diffusion: $P_e = 2$.



(b) Streamline diffusion: $P_e = 200$.



(c) Artificial diffusion: $P_e = 2$.



(d) Artificial diffusion: $P_e = 200$.

Figure 4.1: Solutions and contour plots for $\delta = 0.4$ and $N = 16$.

[4] W. Hackbusch. *Multi-grid Methods and Applications*. Springer-Verlag, New York, 1980.

[5] C. Johnson. *Numerical Solutions of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge, 1987.

[6] K.W. Morton. *Numerical Solution of Convection-Diffusion Problems*. Chapman and Hall, London, 1996.

[7] H.-G. Roos. Necessary convergence conditions for upwind schemes in the two-dimensional case. *Int. J. Numer. Methods Engrg*, 21:1459–1469, 1985.