Scientific Computing
R. Stepleman et al. (eds.)
© IMACS/North-Holland Publishing Company, 1983

DEFECT CORRECTION FOR THE SOLUTION OF A SINGULAR PERTURBATION PROBLEM

P.W. Hemker and P.M. de Zeeuw

Mathematisch Centrum, Amsterdam
The Netherlands

A method is described for the accurate discretization of a differential equation in
which the highest derivative is multiplied by a small parameter. It is well known that
for such singular perturbation problems with a strongly asymmetric differential opera-
tor almost all discretizations are either unstable or inaccurate or direction depen-
dent. By the combination, in an iterative process, of an inaccurate stable and an
accurate unstable scheme we obtain an accurate stable solution, without adapting the
scheme to the flow direction. In fact, two approximate solutions are obtained, that
– uniformly in ε – are both $O(h^2)$ accurate in the smooth part of the solution. The
difference between both solutions can be used for the detection of the unsmooth parts.

1. INTRODUCTION

An iterative process is used to obtain the accu-
rate solution of a singular perturbation problem.
As a model problem we use the convection diffu-
sion equation

(1a) $-\varepsilon\Delta u + \vec{a}.\nabla u = f,$

on a bounded domain $\Omega \subset \mathbb{R}^2$, with either Dirichlet
or natural boundary conditions

(1b) $u = g$ on $\partial\Omega_1$,

(1c) $\vec{n}\varepsilon\nabla u = h$ on $\partial\Omega_2$,

where $\vec{n}$ is the outward normal on the boundary
$\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$.

The problem is written in symbolic form as

(2) $L_\varepsilon u = f.$

It is well known that for such a problem with a
strongly asymmetric differential operator, the
usual discretizations are either unstable (the
usual symmetric discretization methods: central
differences, finite elements or Bubnov-Galerkin
methods) or inaccurate (artificial viscosity) or
direction dependent (various streamline-upwind
or Petrov-Galerkin discretizations). By the
combination, in an iterative process, of an in-
accurate stable and an accurate unstable scheme,
we construct a solution which is accurate,
uniformly for all ε, without adapting the scheme
to the subcharacteristic directions (flow-direc-
tions) in the problem.

In the iterative process two standard discreti-
zations of (2) are used:

(3) $L_{\varepsilon,h} u_h = f_h,$

a standard accurate discretization (e.g. standard
central differences or a finite element discreti-
zation with piecewise linear test and trial
functions on a regular triangularization), which
is unstable for $\varepsilon \ll h$, and an artificial diffusion

(artificial viscosity) discretization

(4) $L_{\alpha,h} u_h = f_h,$

which is the same as (3) but with $\alpha = \varepsilon + O(h)$,
whence (4) is stable. It is well known [6] that
both discretizations yield bad results for small
ratios ε/h.

We combine the discretizations (3) and (4) in an
iterative process of defect correction type. In
case of a linear problem an elementary defect
correction process generates a sequence of appro-
ximate solutions $u^{(i)}$, i = 1,2,3,..... , by the
iteration (iterative refinement)

        $u^{(0)} = 0,$

(5) $\tilde{L}u^{(i+1)} = f - Lu^{(i)} + \tilde{L}u^{(i)}.$

If $\lim_{i\to\infty} u^{(i)} = u^*$ and $\tilde{L}$ is injective, then u^*
is the solution of the "target equation"

        $Lu = f.$

$\tilde{L}$ is usually some approximation to L, for which
the equation (5) is readily solved.

For our problem we study a "mixed" defect correc-
tion iteration with two target operators and two
approximate operators. They are combined as:

        $u^{(0)} = 0,$

(6a) $\tilde{L}_1 u^{(i+\frac{1}{2})} = f_1 - L_1 u^{(i)} + \tilde{L}_1 u^{(i)},$

(6b) $\tilde{L}_2 u^{(i+1)} = f_2 - L_2 u^{(i+\frac{1}{2})} + \tilde{L}_2 u^{(i+\frac{1}{2})}$

If this iteration converges, we obtain two "sol-
utions" $u^A = \lim_{i\to\infty} u^{(i)}$ and $u^B = \lim_{i\to\infty} u^{(i+\frac{1}{2})}$.
With $f_1 = f_2 =: f$, obviously u^A is characterized
by

(7) $[\tilde{L}_2 - (L_2 - \tilde{L}_2)\tilde{L}_1^{-1}(L_1 - \tilde{L}_1)]u^A = [I - (L_2 - \tilde{L}_2)\tilde{L}_1^{-1}]f$

and u^B is given by a similar equation.

2. THE DISCRETIZATION

To obtain an approximate solution for (2), we use the process (6) with $f_1 = f_2 = f_h$ and with the operators

$$(8)\quad L_1 = L_{\varepsilon,h}, \quad \tilde{L}_1 = L_{\alpha,h},$$
$$L_2 = L_{\alpha,h}, \quad \tilde{L}_2 = 2\,\mathrm{diag}\,(L_{\alpha,h}) =: D.$$

In this way the second iteration step (6b) is a damped Jacobi-relaxion-step for the solution of (4). The first step (6a) is an order improving defect correction step towards the solution of (3), where the defect defining operator is simply given by

$$(9)\quad E := \tilde{L}_1 - L_1 = (\alpha-\varepsilon)\Delta_h.$$

The two solutions u_h^A and u_h^B are now characterized by

$$(10)\quad (L_{\varepsilon,h} + L_{\alpha,h}D^{-1}E)u_h^A = f_h,$$
$$(11)\quad (L_{\varepsilon,h} + ED^{-1}L_{\alpha,h})u_h^B = (I + ED^{-1})f_h.$$

By using local mode analysis [4] we can show that, for the smooth parts of the solution, u_h^A and u_h^B are both accurate of $O(h^2)$. Further, the operators $L_{\varepsilon,h} + L_{\alpha,h}D^{-1}E$ and $L_{\varepsilon,h} + ED^{-1}L_{\alpha,h}$ are stable and the numerical boundary layers extend over a region of only $O(h)$ in the boundary layer region [5].

Remark.

It is *not* possible to find an accurate approximation for our problem (1) by application of the simple defect correction (5) alone. If we apply (5) with $L = L_{\varepsilon,h}$ and $\tilde{L} = L_{\alpha,h}$, for $\varepsilon \ll h$ iteration would not converge to a sensible solution because of the instability of $L_{\varepsilon,h}$. It would converge to the unwanted solution of the "target problem" (3). Theoretically, already a single step (or a few steps) of (5) result in a 2nd order accurate method cf. [2] Satz 2.2 , from which follows

$$\| u_h - R_h u\| \le Ch^2\| u''\| .$$

Here $R_h u$ denotes the restriction of the true solution u to the mesh. However this is not a useful errorbound in our case, where $\| u''\|$ may be very large. In fact, for small ε, we do not find the $O(h^2)$ convergence in practice for any reasonable value of h. This is shown in table 1.

$\varepsilon = 1$	$u_h^{(\frac{1}{2})}$	u_h^D	u_h^A	u_h^B
h = 1/8	0.630	0.0740	0.0780	0.0693
ratio	*2.47*	*3.65*	*3.64*	*3.46*
h = 1/16	0.0255	0.0203	0.0214	0.0201
ratio	*1.71*	*4.02*	*4.01*	*3.89*
h = 1/32	0.0149	0.00505	0.00533	0.00516

Table 1a.

$\varepsilon = 10^{-6}$	$u_h^{(\frac{1}{2})}$	u_h^D	u_h^A	u_h^B
h = 1/8	0.790	0.634	0.608	0.459
ratio	*1.37*	*1.76*	*3.82*	*3.48*
h = 1/16	0.578	0.360	0.159	0.132
ratio	*1.52*	*2.08*	*4.75*	*4.54*
h = 1/32	0.380	0.173	0.0335	0.0291

Table 1b.

Table 1 shows the error in the smooth part of the solution in the max-norm.
$u_h^{(\frac{1}{2})}$: the solution after one step (6a);
u_h^D : the solution after two steps (6a);
u_h^A, u_h^B obtained by iteration of (6).
The problem: $\varepsilon\Delta u + u_x = f$ on the unit square; with the Dirichlet boundary data and the data f such that
$u(x,y) = \sin(\pi x)\sin(\pi y) + \cos(\pi x)\cos(3\pi y) + (\exp(-x/\varepsilon) - \exp(-1/\varepsilon))/(1-\exp(-1/\varepsilon)).$

3. RELATION WITH MULTIGRID TECHNIQUES

The algorithm in section 2 can be considered independent of multigrid techniques. However, it is related to previous work done on multigrid methods.

The use of defect correction in combination with a multigrid algorithm is already mentioned by Brandt[1] and is theoretically studied by Hackbusch [2,3]. They consider a multigrid algorithm which, in the elementary form of a two-level algorithm, can be described by a process (6) as well. Then, (6a) is a "coarse grid correction" which is now written in the form

$$u^{(i+1)} = u^{(i)} + \tilde{L}_1^{-1}(f_1 - L_1 u^{(i)}),$$

because $\tilde{L}_1^{-1}$ is of deficient rank:

$$\tilde{L}_1^{-1} = P_{h,2h}L_{\alpha,2h}^{-1}R_{2h,h},$$

where $P_{h,2h}$ and $R_{2h,h}$ are the grid transfer operators between the fine and the coarse grid; (6b) denotes a sequence of relaxation sweeps. In the standard two-level algorithm both target operators are the same: $L_1 = L_2$. In combination with defect correction (non-standard), the operator L_1 corresponds to a more accurate discretization then L_2.

In the approach we discuss here, we differ from [1,2,3] by using a full rank operator $\tilde{L}_1 = L_{\alpha,h}$. We use the multigrid algorithm only to solve efficiently the equation (6a).

4. THE DIFFERENCE BETWEEN u_h^A AND u_h^B

Considering the difference $u_h^B - u_h^A$ of the solutions of (10) and (11), we find the following

Lemma

(12) $\quad u_h^B - u_h^A = D^{-1}Eu_h^A = \frac{\alpha-\varepsilon}{2} \text{diag}^{-1}(L_{\alpha,h})\Lambda_h u_h^A.$

Proof

Because of (6b) we have

(13) $\quad D(u_h^B - u_h^A) = L_{\alpha,h}u_h^B - f_h ,$

and hence, by (11),

(14) $\quad L_{\varepsilon,h}u_h^B - f_h = E D^{-1}(f_h - L_{\alpha,h}u_h^B) =$

$\qquad = (L_{\varepsilon,h} - L_{\alpha,h})(u_h^B - u_h^A) ,$

or

(15) $\quad L_{\alpha,h}u_h^B - f_h = E u_h^A .$

From (13) and (15) the lemma follows immediately.

From (12) we see that the difference $u_h^B - u_h^A$ is large for large $\Lambda_h u_h^A$. In the regions where $u_h^B - u_h^A$ is large, the 2nd order differences of u_h^A are large. Here the solution cannot accurately be represented. Thus, we may use the difference $u_h^B - u_h^A$ as an indicator where the mesh should be refined for a better approximation of the solution.

To study the local behaviour of the difference $u_h^B - u_h^A$ in more detail, we resort to the local mode analysis [1]. We consider the problem (1a) on $\mathbb{R}^2$ and form the discrete operator $L_{\varepsilon,h}$ over all $\mathbb{R}^2$. As a forcing function we take the "mode"

$\quad f_h(jh) = e^{ijh\omega}; \quad j\in\mathbb{Z}^2 ; \quad h\in\mathbb{R}_+^2 ;$
$\qquad\qquad Re(h\omega) \in [-\pi,+\pi]^2.$

Now the solutions u_h^A and u_h^B are studied by Fourier analysis. We note that the Fourier transform is a norm-preserving bijection between the $\ell_2(\mathbb{Z}^2)$ and the $L_2([-\pi,+\pi]^2)$ functions, i.e.

$\qquad \|u\| = \|\hat{u}\| ,$

where $\hat{u}$ is the Fourier transform of u.
From (12) we derive

(16) $\quad \widehat{u_h^B-u_h^A} = FT(\frac{1}{2}(\alpha-\varepsilon)\text{diag}^{-1}(L_{\alpha,h})\Lambda_h u_h^A) =$

$\quad = \frac{\alpha-\varepsilon}{2\alpha}(\sin^2(\omega_1 h/2)+\sin^2(\omega_2 h/2))\hat{u}_h^A =: \frac{\alpha-\varepsilon}{2\alpha}S^2 \hat{u}_h^A,$

and from (10) we find

(17) $\quad \hat{f}_h/\hat{u}_h^A = \frac{-4\varepsilon}{h^2}S^2(1+\frac{\alpha-\varepsilon}{2\varepsilon}S^2) + \frac{4i}{h} T(1+\frac{\alpha-\varepsilon}{2\alpha}S^2),$

where $4iT/h = \hat{L}_{0,h}(\omega)$ is the characteristic trigonometric polynomial for the reduced difference operator.

In previous work [4,5] we derived from (17) that u_h^A is bounded, uniform in ε, Moreover, u_h^A is $\mathcal{O}(h^2)$ accurate in the smooth parts of the solution. We also showed that, for small ε, at boundary or interior layers the error in u_h^A may be $\mathcal{O}(1)$, uniformly in ε. In these regions (where the discrete solution is not able to represent the true solution anyway) the numerical approximation may show oscillations. However, the discretization (10) is asymptotically stable and the critical regions near the layers have only $\mathcal{O}(h)$ width.

To see the effect of $h \to 0$ on the approximation, we study separately the "low" and the "high" frequencies in the true solution. The frequencies are called "high" or "low" with reference to the mesh used (size h). For the "low" frequencies in the solution we consider a fixed ω and let $h \to 0$. For the "high" frequencies we consider ωh fixed and let $h\to 0$. Regions where the solution is smooth are characterized by low frequencies. Regions where the mesh is too coarse to represent a solution properly, are dominated by high frequencies.

Now we use (16) to see how $u_h^B - u_h^A$ behaves in the different regions. For the low frequencies $S^2 = \sin^2(\omega_1 h/2) + \sin^2(\omega_2 h/2) = \mathcal{O}(h^2)$ and we find

$\| u_h^B - u_h^A \| = \frac{\alpha-\varepsilon}{2\alpha} S^2 \| u_h^A \| \le C \frac{\alpha-\varepsilon}{\alpha} h^2 \| u_h^A \|$

$\qquad \approx C \frac{h}{h+\varepsilon} h^2 \| u_h^A \| .$

Because u_h^A is bounded, uniform in ε, we conclude that in the smooth part of the solution we have

$\| u_h^B - u_h^A \| \le C h^2 ,$

with C independent of ε.
For the high frequencies ωh is fixed (and hence also S^2) and we find

$\| u_h^B - u_h^A \| = \frac{\alpha-\varepsilon}{2\alpha} S^2 \| u_h^A \| \le \frac{\alpha-\varepsilon}{\alpha} \| u_h^A \|$

$\qquad \le C h/(h+\varepsilon).$

Hence, for $h<\varepsilon$, ε fixed, we find

$\| u_h^B - u_h^A \| = \mathcal{O}(h)$

and for $\varepsilon<h$ we find

$\| u_h^B - u_h^A \| = \mathcal{O}(1)$

for the high frequencies.
Hence we find that in the smooth part of the solution (low frequencies) the difference $u_h^B - u_h^A$ is small, whereas it may be large in those regions of transition layers, where the mesh is not fit to represent the true solution. In this way the difference $u_h^B - u_h^A$ behaves similar to the error in the approximation $u_h^A - R_h u$.

5. EXAMPLES

As a first example to show the behaviour of u_h^A and $u_h^B - u_h^A$, we use the same equation as Hughes and Brooks [6],

$\quad -\varepsilon\Delta u + \cos(\theta) u_x + \sin(\theta) u_y = 0,$

on the unit square, with Dirichlet boundary conditions at the inflow boundary and $\varepsilon = 10^{-6}$.
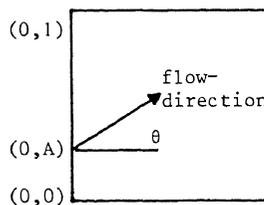
Inflow boundary conditions

$u(x,0) = 1, \quad 0 \le x \le 1,$
$u(0,y) = 1, \quad 0 \le y \le A,$
$u(0,y) = 0, \quad 1 \ge y > A.$

$A = 3/16.$
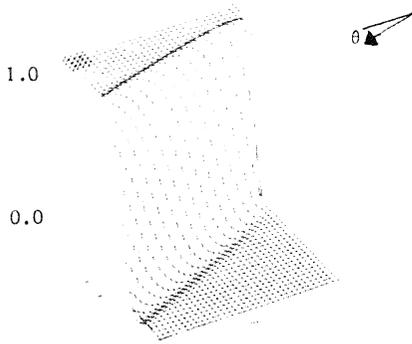
Fig. 1a. Example 1. Neumann BCs, $\theta = 22.5°$, u_h^A.
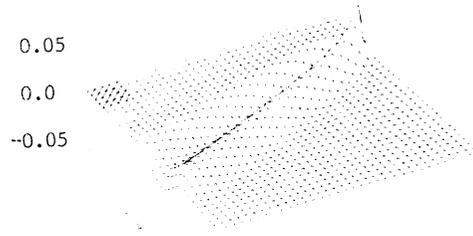
Fig. 1b. $u_h^B - u_h^A$. ▨
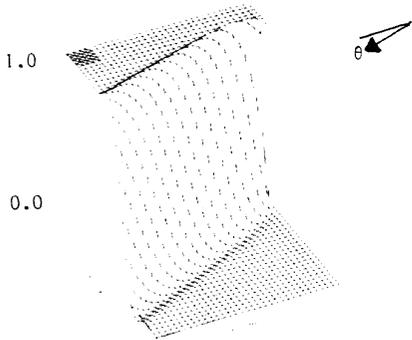
Fig. 2a. Example 1. Neumann BCs, $\theta = 22.5°$, u_h^A.
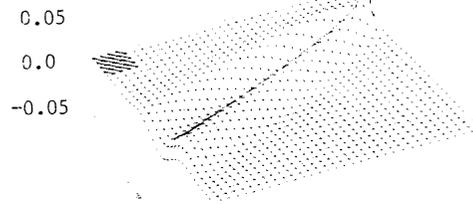
Fig. 2b. $u_h^B - u_h^A$. ▨
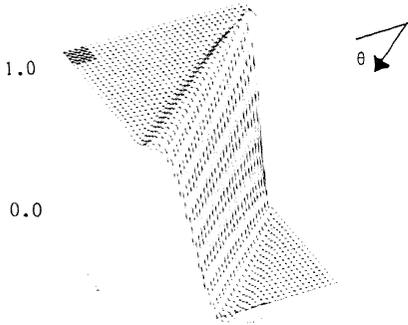
Fig. 3a. Example 1. Neumann BCs, $\theta = 45°$, u_h^A.
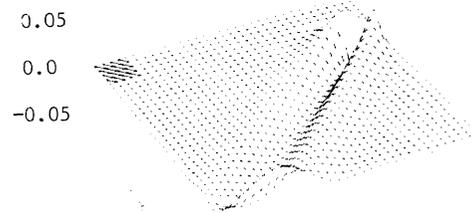
Fig. 3b. $u_h^B - u_h^A$.
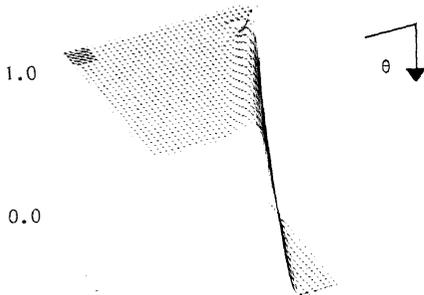
Fig. 4a. Example 1. Neumann BCs, $\theta = 67.5°$, u_h^A.
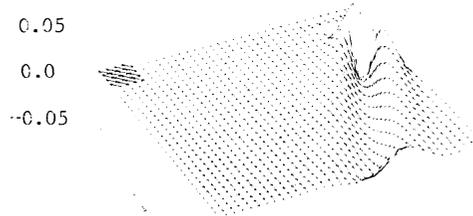
Fig. 4b. $u_h^B - u_h^A$.
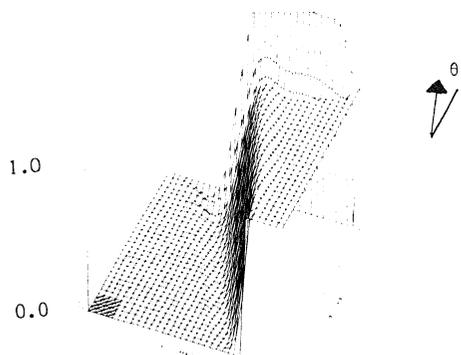
Fig. 5a. Example 1. Dirichlet BCs, $\theta = 22.5°$, u_h^A.
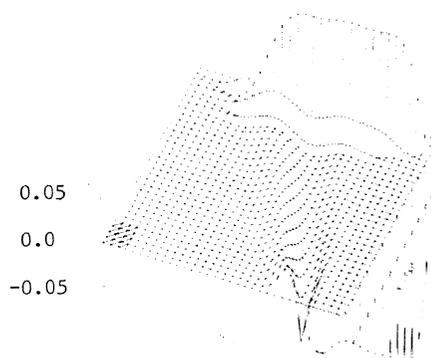
Fig. 5b. $u_h^B - u_h^A$.
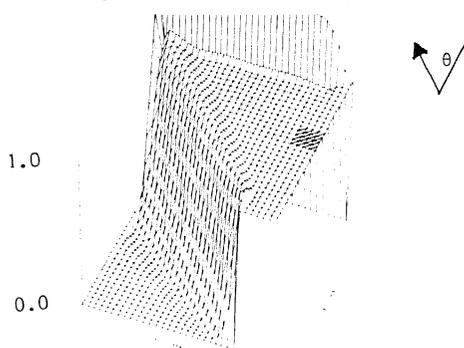
Fig. 6a. Example 1. Dirichlet BCs, $\theta = 45°$, u_h^A.
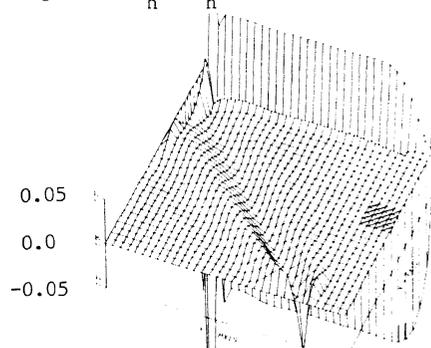
Fig. 6b. $u_h^B - u_h^A$.
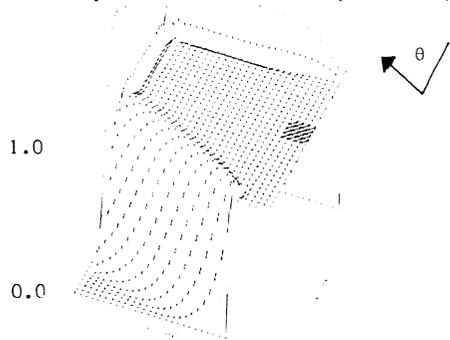
Fig. 7a. Example 1. Dirichlet BCs, $\theta = 67.5°$, u_h^A.
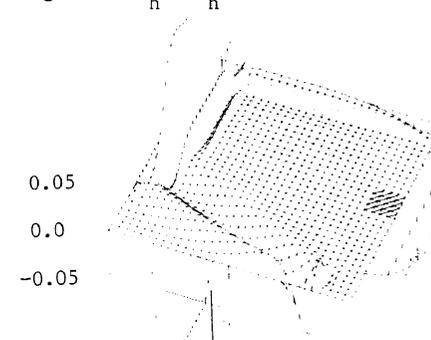
Fig. 7b. $u_h^B - u_h^A$.
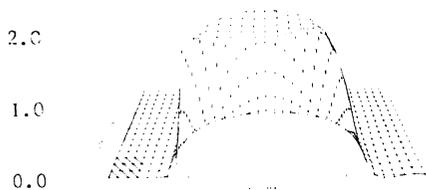
Fig. 8a. Example 2. u_h^A.
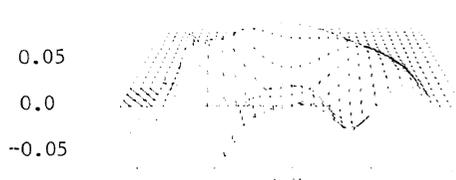
Fig. 8b. Example 2. $u_h^B - u_h^A$.

At the outflow boundary either natural or homo-
geneous Dirichlet boundary conditions are used.
For $L_{\varepsilon,h}$ we use the usual finite element discre-
tization with piecewise linear functions on a
regular triangularization.
Note: the way the squares are triangularized
(◺ or ◿) makes no essential difference in the
results. If the triangle division follows the
internal layer, the numerical results are
-of course- slightly better. In the figures
the triangularization is indicated.

In common with [6] we use as flow directions
θ = 22.5°, 45° and 67.5°.
In the figures (1)-(7) we show the numerical
solution u_h^A and the difference $u_h^B - u_h^A$ for the
different cases.

As a second example we use the variable
coefficient problem :

$$-\varepsilon \Delta u + \vec{a} . \nabla u = 0 \quad \text{on } [-1,+1] \times [0,1],$$
$$\vec{a} = (\ y(1-x^2), \ -x(1-y^2))^T,$$

with the Dirichlet boundary conditions

$u(x,y) = 1 + \tanh(10+20x)$, $y=0$, $-1 \le x \le 0$,
$u(x,y) = \quad\quad 0 \quad\quad$, $y=1$, $-1 \le x \le 1$,
$u(x,y) = \quad\quad 0 \quad\quad$, $x=1$, $0 \le y \le 1$,
$u(x,y) = \quad\quad 0 \quad\quad$, $x=-1$, $0 < y \le 1$;

and homogeneous Neumann boundary conditions at
the outflow boundary: $y=0$, $0 < x < 1$.
Asymptotically for $\varepsilon \to 0$, the true solution is
constant over the subcharacteristics and the
outflow profile is the mirror immage of the
inflow profile.
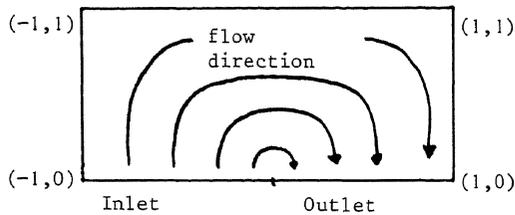
In the figures (8a) and (8b) we show the
numerical solution u_h^A and the difference
$u_h^B - u_h^A$.

REFERENCES

[1] A.Brandt, Numerical stability and fast sol-
    utions to boundary problems.
    In: Boundary and interior layers - Computa-
    tional and asymptotic methods.(J.J.H.Miller
    ed.) Boole Press,Dublin,1980.

[2] W.Hackbusch, Bemerkungen zur iterierten
    Defektkorrektur und zu ihrer Kombination mit
    Mehrgitterverfahren.
    Rev.Roum.Math.Pures Appl.26(1981)1319-1329.

[3] W.Hackbusch, On multigrid iterations with
    defect correction.
    In: Multigrid Methods (W.Hackbusch and

U.Trottenber,eds) LNM, Springer, 1982.

[4] P.W.Hemker, An accurate method without dir-
    ectional bias for the numerical solution of
    a 2-D elliptic singular perturbation prob-
    lem. In: Singular perturbations and
    applications (W.Eckhaus and E.M. de Jager
    eds) LNM,Springer,1982.

[5] P.W.Hemker, Mixed defect correction iter-
    ation for the accurate solution of the
    convection diffusion equation.
    In: Multigrid Methods (W.Hackbusch and
    U.Trottenberg, eds) LNM, Springer,1982.

[6] T.J.R.Hughes and A.Brooks, A multi-
    dimensional upwind scheme with no crosswind
    diffusion.
    In: Finite Element Methods for Convection
    Dominated Flows (T.J.R.Hughes ed.), ASME,
    New York, 1979.

[7] P.Gresho and R.L.Lee, Don't suppress the
    wiggles - They're telling you something.
    In: Finite Element Methods for Convection
    Dominated Flows (T.J.R.Hughes ed.), ASME,
    New York, 1979.