SIAM J. SCI. COMPUT.
Vol. 14, No. 3, pp. 607–626, May 1993
© 1993 Society for Industrial and Applied Mathematics
008

ACCELERATED MULTIGRID CONVERGENCE AND HIGH-REYNOLDS RECIRCULATING FLOWS*

A. BRANDT† AND I. YAVNEH‡

Abstract. Techniques are developed for accelerating multigrid convergence in general, and for advection-diffusion and incompressible flow problems with small viscosity in particular. It is shown by analysis that the slowing down of convergence is due mainly to poor coarse-grid correction to certain error components, and means for dealing with this problem are suggested, analyzed, and tested by numerical experiments, showing very significant improvement in convergence rates at little cost.

Key words. multigrid, incompressible flow, acceleration

AMS(MOS) subject classifications. 65N20, 76D05

1. Introduction.

Classical multigrid methods were originally developed for elliptic partial differential equations and systems. For such problems these methods have proved to be extremely efficient, enabling solution to the level of discretization errors in just a few minimal work units, so that the total work invested in the solution grows linearly with the number of variables, and usually at most several dozen operations per variable are required. When applied to nonelliptic and singular perturbation problems such as high-Reynolds flows, however, performance seems to deteriorate significantly, due in part to the fact that the solutions become more complex (e.g., boundary layers and characteristic directions along which high-frequency data may propagate to large distances). But poor multigrid behavior is exhibited even in simple problems with smooth solutions when the partial differential operator is nonelliptic (or has a nonelliptic component).

To learn how to treat the various troubles that appear in flow problems, it is necessary to distinguish them. One of the main distinctions that must be made is between *entering flows*, in which the flow enters through some boundary and follows a well-defined general orientation, and *recirculating* flows. In the former, relaxation can be made to resolve smooth components, and thus act (in part) as a *solver* and not just as a *smoother* in the multigrid solution process. Highly efficient multigrid solutions to problems of this type are demonstrated in [7]. The purpose of the present study is to show how to improve performance greatly in recirculating flows and other problems, in which relaxation significantly resolves only nonsmooth error components.

The main tool used in analysis and prediction of the performance of multigrid algorithms is local mode analysis (see, e.g., [2]). For elliptic partial differential systems it was shown in [4] that, under reasonable assumptions, the predicted performance can indeed be obtained for general domains by adding some processing, for negligible cost, at and near the boundaries. Hence, boundary effects may be neglected in the basic

*Received by the editors March 11, 1991; accepted for publication (in revised form) April 20, 1992. This research was supported in part by the Air Force Office of Scientific Research, United States Air Force, under grant AFOSR-91-0156, by the National Science Foundation under grant NSF DMS-9015259, and by the Department of Energy under contracts W-7405-ENG-36 and KC-07-01-01, while the second author was visiting the Center for Nonlinear Studies in the Los Alamos National Laboratory.

†Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel.

‡Oceanography Department, National Center for Atmospheric Research, P.O. Box 3000, Boulder, Colorado 80307-3000. This research was partly performed while the author was at the Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel.

analysis of multigrid solvers for elliptic problems. This is not the case in problems that are nonelliptic or have nonelliptic components, such as incompressible flows at high-Reynolds numbers. In such problems, high-frequency boundary data may propagate far into the domain, and it may be necessary to include boundary effects in the analysis (see [7]). However, when dealing with shear-driven recirculating flows we are once again in a position that is similar to the elliptic case with respect to boundary effects, since cross-stream behavior is determined by the (elliptic) viscosity terms no matter how small they may be (see [6]). For such flows the infinite-space local mode analysis is again useful.

Slow convergence of multigrid cycles can generally be traced to (at least) one of two major causes: poor smoothing or poor coarse-grid approximation to the fine-grid problem. Smoothing, which is reducing the amplitude of high-frequency errors that cannot be treated on the next coarser grid, is dealt with extensively in most publications on multigrid solvers for incompressible flows (see, e.g., [2] and [5]). Here we address the problem of poor coarse-grid corrections, which may require specialized methods that depend on the cause of the problem. When the poor approximation is caused by a local factor, say, some singularity in the boundary, the proper course is usually to employ local methods such as extra relaxation in the vicinity of the singularity (see [1] and [4]). But sometimes, in particular in high-Reynolds flows, the coarse grid fails to approximate the fine-grid problem well enough for certain components *throughout the domain.* Frequently, the best course of action then is to simply disregard this seemingly poor behavior, since those components, which are poorly approximated by the coarse grid, do not in turn approximate the differential solution well (see [2] and [3]). If, however, one is interested in obtaining good *algebraic* convergence and not just rapid convergence to a good approximation of the *differential* solution, special techniques must be employed to accelerate convergence.

Consider as our model problem the constant-coefficient advection-diffusion equation:

$$L_{ad}u = -\epsilon\triangle u + au_x = 0\,, \tag{1}$$

where $\triangle$ is the Laplacian operator and ϵ is a positive constant. Suppose that this equation is discretized by some finite difference scheme L^h_{ad} of order q on a uniform grid of mesh-size h whose orientation is general (and therefore choosing x to be the characteristic direction results in no loss of generality). The resulting discrete set of equations is solved by starting with some initial approximation to the discrete solution on the fine grid h (grid with mesh-size h) and iterating with the usual multigrid cycles. Soon one finds that (for vanishing ϵ) the residual norms are reduced at best by an amplification factor of $1 - 0.5^q$ by each cycle, even if the number of relaxation sweeps per level and the cycle index γ (defined below) are chosen to be quite large. The reason for this slow-down has been shown to be poor approximation of *smooth characteristic components* by the coarse grids. This property has already been explained in [3]. We return to it here, and in another context in [7], and present methods for treating the problem.

As usual, when researching the advection-diffusion equation, our object is to learn how to treat the Navier–Stokes equations. Since these exhibit similar behavior, we conclude that the problem of poor coarse-grid correction studied here is the main cause for the poor convergence rates of flow problems as well. Hence, although the analyses below are all done for the advection-diffusion equation, numerical experiments also include the incompressible flow equations, and indeed the behavior is seen to be influenced similarly by the methods proposed.

In §§2 and 3 the two-level and multilevel cycles are analyzed, and it is shown why the usual multigrid cycles exhibit poor convergence rates. In §4 it is shown by analysis and numerical examples that significant acceleration of convergence can be obtained at virtually no cost by overweighting some of the residuals that are transferred to the next coarser grid. In §5 a method of employing defect corrections within the multigrid cycle is introduced, and it is shown by analysis and numerical experiments that this method can be combined in various ways with residual overweighting to greatly improve the multigrid cycle's performance. The ultimate modified W cycle that is developed enables an asymptotic reduction of the error by a factor of nine per cycle in its two-level version for the first-order discretized advection-diffusion and incompressible Navier–Stokes equations, rather than the factor of two that is yielded by the usual two-level cycle. This performance is almost matched in the multilevel cycle as well, and here the improvement is still more dramatic, since the usual cycle's performance is shown to deteriorate rapidly as the number of levels grows. Conclusions and remarks are given in §6.

The accelerated multigrid analyses and methods developed in this article can be straightforwardly generalized to other cases in which multigrid convergence deteriorates due to poor coarse-grid corrections for some smooth components.

2. Two-level infinite-space local mode analysis. In order to have a clear and quantitative understanding of the slowing down of multigrid convergence due to poor coarse-grid correction, we will analyze here a multigrid cycle leaving out some irrelevant aspects: we will treat only smooth components, and hence neglect the effect of intergrid transfers, and employ the first differential approximation (FDA; see [11], [3], and [2, §7.5]) to the difference equations. We will also assume for simplicity that the diffusion coefficient ϵ tends to zero, which is indeed the case for which the multigrid performance is usually the worst. The FDA approximation to a qth-order discretization of L_{ad} then has the form

$$L_{ad}^h = -\epsilon_h a T_{q+1}(\partial_x, \partial_y) + a\partial_x \,, \tag{2}$$

where the first term on the right-hand side represents the first truncated term in the discretization of L_{ad}. T_{q+1} is a polynomial of degree $q+1$ of the form

$$T_{q+1}(\partial_x, \partial_y) = \sum_{j=0}^{q+1} a_j \, \partial_x^j \, \partial_y^{q+1-j}, \tag{3}$$

and $\epsilon_h = O(h^q)$. ϵ_h and the coefficients a_j are determined by the particular discretization and angle between the characteristic direction and the grid. For example, in the case of the usual first-order upstream discretization, the FDA is

$$\begin{aligned} L_{ad}^h = -\tfrac{1}{2} h a [\, &(\cos^3\phi + \sin^3\phi)\partial_{xx} + \sin 2\phi(\sin\phi - \cos\phi)\partial_{xy} \\ &+ \tfrac{1}{2}\sin 2\phi(\sin\phi + \cos\phi)\partial_{yy}\,] + a\partial_x \,, \end{aligned}$$

where ϕ is the angle that the characteristic direction x forms with the grid. (This is obtained by writing the artificial viscosity of upstream discretization in the coordinates of the grid, and transforming to the characteristic coordinates x and y.)

Consider an error component on the fine grid h given by

$$v^h = e^{i(\omega_1 x + \omega_2 y)} \,. \tag{4}$$

Generalization to higher dimensions is straightforward, and the following analyses apply directly. The residual due to this error is

$$r^h = -L^h_{ad} v^h = -\hat{L}^h_{ad}(\omega_1, \omega_2) \cdot e^{i(\omega_1 x + \omega_2 y)}, \tag{5}$$

where $\hat{L}^h_{ad}$ is called the *symbol* of L^h_{ad} (see also [2]). The correction problem solved on the coarse grid H is

$$L^H_{ad} v^H = r^h, \tag{6}$$

yielding

$$v^H = -\frac{\hat{L}^h_{ad}}{\hat{L}^H_{ad}} v^h. \tag{7}$$

Denoting

$$C(\omega_1, \omega_2) = \frac{\hat{L}^h_{ad}(\omega_1, \omega_2)}{\hat{L}^H_{ad}(\omega_1, \omega_2)},$$

we obtain that the two-level amplification factor ν of the fine-grid error, defined as the ratio of the new fine-grid error (after the correction v^H has been added) to the original error v^h, is given by

$$\nu(\omega_1, \omega_2) = \frac{v^h + v^H}{v^h} = 1 - C(\omega_1, \omega_2). \tag{8}$$

From (2), (3), and (8), for a qth-order discretization of the advection-diffusion operator, we obtain

$$\nu(\omega_1, \omega_2) = 1 - \frac{\epsilon_h T_{q+1}(i\omega_1, i\omega_2) + i\omega_1}{\epsilon_H T_{q+1}(i\omega_1, i\omega_2) + i\omega_1}. \tag{9}$$

For most smooth components (and small mesh-sizes) this amplification factor is close to zero, and multigrid convergence can hence be expected to be essentially as good as allowed by the smoothing rate and the number of relaxation sweeps per level. But consider smooth *characteristic error components*, which are smooth (and therefore nearly unaffected by any local-type relaxation), but much smoother in the characteristic direction x than in the cross-characteristic direction y. For such components, $\omega_1/\omega_2 \ll 1$, and their amplification factor is therefore much larger. In the limit of components which only vary in the cross-characteristic direction, we obtain that the two-level *convergence factor* ν_{tl}, which is defined as the maximal (in absolute value) ν over all the frequencies defined on the grid, is given by

$$\nu_{tl} = \nu(0, \omega_2) = 1 - \frac{\epsilon_h}{\epsilon_H} \approx 1 - 0.5^q. \tag{10}$$

Asymptotically the poorly corrected characteristic components become dominant, and the error norm is reduced at best by a factor of $1 - 0.5^q$, even with two-level cycles.

3. Multilevel local mode analysis. The multigrid *cycle index* γ is defined as the number of times a correction from the next coarser grid is taken by each

intermediate level in calculating its solution, before that solution is used to correct the error on the next finer grid. Thus $\gamma = 1$ is the index of the so-called V cycle, and $\gamma = 2$ is the index of the W cycle. Actual multigrid cycles need to employ a fairly small cycle index γ if the total work is to grow only linearly with the number of variables: in d dimensions, γ must be smaller than 2^d. Let us therefore determine by analysis the expected multilevel amplification factor of an error component whose two-level amplification factor $1 - C$ is assumed to be independent of the mesh-size. This simplifying assumption is only valid for the extreme cases where, in (9), ω_1 vanishes, in which case C equals $\epsilon_h/\epsilon_{2h} \approx 0.5^q$, or where T_{q+1} vanishes, in which case C equals one. This is appropriate here since these extreme cases determine the worst-case amplification factor. Moreover, this analysis of the multigrid cycle is not restricted to the advection-diffusion equation.

Consider a multigrid cycle that is implemented on $n+1$ levels numbered $0, \ldots, n$, with n denoting the finest grid. For each error component there corresponds a "solution level" i on which it is sufficiently nonsmooth so as to be eliminated efficiently by relaxation (or perhaps direct solution in the case $i = 0$). To simplify, we assume that every error component is eliminated completely on its solution level, but that on all finer grids it is unaffected by relaxation. Let us denote by $\nu_\gamma^{n-i}(C)$ the level-n amplification factor of an error component, whose two-level amplification factor is $1 - C$ and whose solution level is i, in a multigrid cycle with cycle index γ. Let $k = n - i$. Under the assumptions above, we have for $k = 0$

$$\nu_\gamma^0 = 0\,, \tag{11}$$

and for $k = 1$

$$\nu_\gamma^1 = 1 - C, \tag{12}$$

regardless of γ (the latter is simply the two-level cycle by definition). For $1 < k \leq n$ the residual problem is transferred to the next coarser grid (level $k-1$), and γ multigrid cycles (in which the corresponding amplification factor for that component is ν_γ^{k-1}) are performed. As a result, the error in the solution to the coarse-grid problem is reduced by a factor of $(\nu_\gamma^{k-1})^\gamma$. This approximate correction is now transferred back to the fine grid, but multiplied by C (since even the exact solution to the coarse-grid problem only yields C times the required correction). This yields the following recurrence equation:

$$\nu_\gamma^k = 1 - C \cdot \left[1 - (\nu_\gamma^{k-1})^\gamma\right] = \nu_\gamma^1 + C(\nu_\gamma^{k-1})^\gamma\,, \qquad 1 \leq k \leq n\,. \tag{13}$$

3.1. Properties of the multilevel cycle. In the following discussion we consider only real values of C, since C is real in both extreme cases of the advection-diffusion equation: when the coarse-grid correction is worst, and when it is best. This assumption will greatly simplify the discussion. Also, it is of course unnecessary to consider nonpositive real values for C, since these would imply that the coarse grid does not approximate the corresponding components at all, in which case special measures need to be taken.

Remark 1. The fixed points of (13) are solutions of

$$\nu = 1 - C(1 - \nu^\gamma)\,. \tag{14}$$

$\nu = 1$ is a fixed point for every γ. When $C < 1$ it is the only one in the case of a V cycle ($\gamma = 1$). W cycles ($\gamma = 2$) have a second fixed point at $\nu = (1 - C)/C$.

PROPOSITION 1. *For every* $k \geq 1$,

$$\nu_1^k = 1 - C^k . \tag{15}$$

Proof. By induction: $\nu_1^1 = 1 - C$, and by (13) and the induction hypothesis,

$$\nu_1^k = 1 - C + C(1 - C^{k-1}) = 1 - C^k . \qquad \square$$

Conclusion. The convergence rate of the multigrid V cycle deteriorates rapidly as the number of levels grows, unless C is close to 1. In particular, the V cycle is clearly unsuitable for the advection-diffusion equation, even with first-order discretization.

PROPOSITION 2. *For all* $0 < C < 1$ *and all positive* k *and* γ,

$$0 \leq \nu_\gamma^{k-1} < \nu_\gamma^k < 1. \tag{16}$$

Proof. By induction: $0 < \nu_\gamma^1 = 1 - C < 1$, and by (13) and the induction hypothesis,

$$\nu_\gamma^{k+1} = 1 - C \cdot [1 - (\nu_\gamma^k)^\gamma] < 1 \tag{17}$$

and

$$\nu_\gamma^{k+1} - \nu_\gamma^k = C \cdot \left[(\nu_\gamma^k)^\gamma - (\nu_\gamma^{k-1})^\gamma\right] > 0 . \tag{18}$$

$\square$

PROPOSITION 3. *For all* $0 < C < 1$, ν_2^k *tends to the smaller of the two fixed points* (1 *and* $(1 - C)/C$) *as* k *tends to infinity.*

Proof. For $k \geq 0$, by (16) we have

$$\begin{aligned} \frac{\left|\frac{1-C}{C} - \nu_2^{k+1}\right|}{\left|\frac{1-C}{C} - \nu_2^k\right|} &= \frac{\left|\frac{1-C}{C} - 1 + C \cdot [1 - (\nu_2^k)^2]\right|}{\left|\frac{1-C}{C} - \nu_2^k\right|} \\ &= \frac{\frac{1}{C}\left|(1-C)^2 - (C\nu_2^k)^2\right|}{\frac{1}{C}\left|1 - C - C\nu_2^k\right|} = \left|1 - C(1 - \nu_2^k)\right| < 1 . \end{aligned} \tag{19}$$

The proof follows from this and Proposition 2. $\square$

Conclusion. Proposition 3 implies that ν_2^k tends to one when $0 < C \leq 0.5$. Thus, the convergence factor of the W cycle for the advection-diffusion equation with first-order discretization tends to one as the number of levels tends to infinity.

We reiterate that this analysis assumes that relaxation is of a local type, and therefore has a negligible effect on all components that are smooth on the scale of the grid on which they are relaxed. A different situation may arise in cases such as *entering flow* problems, which are studied in [7]. When relaxation is performed in downstream ordering in such problems, it no longer has a purely local effect. Indeed, for the advection equation with upstream differencing, if relaxation is carried out in downstream ordering, it performs as a *solver* and not just as a *smoother*. Convergence of the advection-diffusion equation is discussed in [8], and much better rates than the ones predicted by our analysis are obtained. But the model problems are precisely of the entering flow type, and at least some of the relaxation sweeps are carried out in downstream ordering. This point, however, which explains a number of phenomena in the numerical results that are obtained there as well as in other publications, and which is the key to further achievements with much more complicated schemes, seems

to have been overlooked. Also, in some flow problems, the actual Reynolds number in the regions of recirculation are very much smaller than the nominal values (because the velocities in these regions are small), and somewhat better performance may then be found (e.g., in [10]).

4. Method of overweighted residuals. A very simple and almost cost-free approach to accelerating convergence is the method of overweighted residuals (OWR). The idea, which is reminiscent of the method of successive overrelaxation (SOR), is to improve the coarse-grid correction to the error in the fine-grid approximation by multiplying the residuals that are transferred to the coarse grid by some constant η between one and two. Clearly, η should not be too large, since those components that usually receive the proper correction are now overcorrected. It is best to determine an optimal overweighting factor η, which will depend on the number of levels employed.

4.1. Two-level single-parameter optimization. Let ξ denote the minimal C over all the frequencies defined on the fine grid (for the advection-diffusion case, $\xi = 0.5^q$). Then the poorest (largest) corresponding error-amplification factor in the usual two-level cycle (with $\eta = 1$) is $1 - \xi$, whereas the best amplification factor is zero. When the residuals are multiplied by the factor η, however, the two extreme error-amplification factors are now given by $1 - \eta\xi$ and $1 - \eta$ (by (6–8)). Due to the monotonicity of these terms, the optimal two-level overweighting factor η_{tl} is obtained when

$$|1 - \eta_{tl}\xi| = |1 - \eta_{tl}|\,, \tag{20}$$

and therefore

$$\eta_{tl} = \frac{2}{1+\xi}\,. \tag{21}$$

The two-level convergence factor ν_{tl} decreases accordingly from $1 - \xi$ to

$$\nu_{tl} = 1 - \eta_{tl}\xi = \frac{1-\xi}{1+\xi}\,. \tag{22}$$

The two-level convergence factor of the advection-diffusion equation (with vanishing diffusion coefficient) employing first-order discretization, for example, improves from 0.5 to 0.33 with $\eta_{tl} = 4/3$.

4.2. Multilevel single-parameter optimization. Propositions 2 and 3 imply that the components that converge most slowly, for a fixed C in (0,1), are those whose solution level is the lowest ($k = n$). When the number of levels is large, the amplification factor of such errors by a W cycle tends to $(1 - C)/C$ when C is between one half and one. For the *overcorrected* components in a cycle with an even cycle index γ, however, we have the following.

PROPOSITION 4. *For all* $1 < C < 2$, *even* γ, *and* $k \geq 1$,

$$|\nu_\gamma^k| \leq |\nu_\gamma^1| = C - 1\,. \tag{23}$$

Proof. By induction: For $k \geq 1$

$$C - 1 - |\nu_\gamma^{k+1}| = C - 1 - |1 - C + C(\nu_\gamma^k)^\gamma|\,, \tag{24}$$

but

$$C - 1 + 1 - C + C(\nu_\gamma^k)^\gamma > 0\,, \tag{25}$$

and alternatively

$$C - 1 - (1 - C + C(\nu_\gamma^k)^\gamma) \geq 2(C-1) - C(C-1)^\gamma > 0\,, \tag{26}$$

where in (26) the first inequality is trivial for $k = 1$ and is due to the induction hypothesis for $k > 1$, and the second is due to the range of C. □

Conclusion. The components that converge most slowly when $1 < C < 2$ are those that are corrected on the *second-finest* level ($k = 1$), and the convergence factor for such overcorrected components is at least as good as the two-level rate when γ is even. This is true in particular for the W cycle ($\gamma = 2$). Hence, the optimal multilevel overweighting factor η_{ml} for the W cycle is obtained when

$$\left|\frac{1 - \eta_{ml}\xi}{\eta_{ml}\xi}\right| = |1 - \eta_{ml}|\,, \tag{27}$$

yielding

$$\eta_{ml} = \xi^{-1/2}\,. \tag{28}$$

This value is valid only when $\eta_{ml} < 2$. For $\xi \leq 0.25$ the multilevel convergence factor must tend to one as the number of levels grows. The W-cycle multilevel convergence factor ν_{ml} of the advection-diffusion equation with vanishing diffusion coefficient and first-order discretization thus improves from 1 to $\sqrt{2} - 1 \approx 0.41$, with the optimal overweighting factor $\eta_{ml} = \sqrt{2}$.

With a finite number of levels, the optimal overweighting factor is reduced somewhat, yielding slightly better convergence rates. It is thus possible to calculate the optimal overweighting factors for each level. These start with 4/3 for the second-coarsest grid and increase until they tend to $\sqrt{2}$ as the grid becomes finer. However, there is not much practical gain in doing such careful optimization, since on very coarse grids the assumptions made in the analysis are rather poor anyway, and also the optimal η quickly tends to $\sqrt{2}$. *One can thus uniformly use the overweighting parameter* $\eta = \sqrt{2}$.

4.3. Two-level multiparameter optimization. In the optimizations described above we have considered only a single cycle, and therefore only a single overweighting parameter η. But since we are dealing with asymptotic convergence rates, it is implied that several cycles are performed. Also, since (at least) W cycles need to be employed, more than one cycle per level is performed. This suggests using several different overweighting factors in order to further reduce the error-amplification factor. This is a special case of *polynomial acceleration*, and the optimal choice of overweighting factors, which is calculated with the aid of *Chebyshev polynomials*, yields *optimal Chebyshev acceleration* (see, e.g., [9]).

Consider a smooth error component for which the ratio between the fine-grid and coarse-grid symbols is C, $\xi \leq C \leq 1$. Suppose that m two-level cycles are performed, where the residuals in the jth cycle, $j = 1, \ldots, m$, are overweighted by a factor of η_j. The factor by which the error is amplified in this process is given by

$$[\nu^m(C)]^m = \prod_{j=1}^{m}(1 - \eta_j C)\,, \tag{29}$$

where $\nu^m(C)$ is the average amplification factor per cycle. Assuming, as above, that the extreme eigenvalues of the two-level operator are 0 and $1-\xi$, the optimal polynomial $\nu_{tl}^m(C)$, i.e., the one for which the maximal absolute value over all C's in the interval is minimized, is given by

$$\nu_{tl}^m(C) = \left[\frac{\mathcal{T}_m\left(\frac{1+\xi-2C}{1-\xi}\right)}{\mathcal{T}_m\left(\frac{1+\xi}{1-\xi}\right)}\right]^{1/m}, \tag{30}$$

where $\mathcal{T}_m$ is the Chebyshev polynomial of degree m, given by

$$\mathcal{T}_m(x) = \begin{cases} \cos(m\cos^{-1}x), & -1 \le x \le 1, \\ \cosh(m\cosh^{-1}x), & 1 < x \end{cases} \tag{31}$$

(see [9] for the general development). By (29) the optimal η_j's are the inverses of the zeros of ν_{tl}^m, which can easily be calculated from (31). For $m = 1$ the optimal overweighting factor is indeed as in (21), and for $m = 2$ the two optimal factors are given by

$$\eta_j = \frac{2}{1+\xi \pm \frac{1-\xi}{\sqrt{2}}}, \qquad j = 1,2. \tag{32}$$

The average amplification factor per cycle of the error is given by

$$\nu_{tl}^m(1) = \left[\mathcal{T}_m\left(\frac{1+\xi}{1-\xi}\right)\right]^{-1/m}. \tag{33}$$

For the advection-diffusion equation with first-order discretization ($\xi = 0.5$), the single-cycle optimization ($m = 1$) thus yields a two-level convergence factor of 0.33 per cycle as shown above. Two-cycle optimization ($m = 2$) yields an average two-level convergence factor of 0.24 with the optimal overweighting factors of 1.079 and 1.745, and the asymptotic (infinite number of cycles) average two-level convergence factor with optimal overweighting factors is 0.17. With second-order discretization ($\xi = 0.25$) the single-cycle optimization yields a two-level convergence factor of 0.60 per cycle with $\eta = 1.60$. Two-cycle optimization yields an average two-level convergence factor of 0.47 with the optimal overweighting factors of 1.123 and 2.779, and the asymptotic average two-level convergence factor with optimal overweighting is 0.33. However, large overweighting factors are unlikely to be useful in practice, if only because the corresponding amplification of nonsmooth error components means that much better smoothing is then required.

An automatic acceleration method, which would not require a priori analysis, is conceivably useful. However, since the ultimate goal is to employ at most one or two cycles per level in the solution process, methods whose usefulness relies on the execution of many cycles are unlikely to be truly efficient.

4.4. γ-cycle optimization. In a multigrid cycle with cycle index γ there are γ overweighting factors to be chosen for the γ visits to the next coarser grid at each level. Consider again a smooth error component for which the ratio between the fine-grid and coarse-grid symbols is C, $\xi \le C \le 1$. Then, if the overweighting factors η_j are chosen independently of the level, the error amplification factor *of γ cycles* on level $i+k$ is given by the recurrence equation

$$(\nu_\gamma^k)^\gamma = \prod_{j=1}^{\gamma} \left[1 - \eta_j C \left(1 - (\nu_\gamma^{k-1})^\gamma\right)\right], \tag{34}$$

with

$$\nu_\gamma^0 = 0. \tag{35}$$

Again a recurrence equation is obtained, for which 1 is a fixed point. The optimization of the η_j's generally needs to be done numerically. The optimal overweighting factors for the W cycle with $\xi = 0.5$ (e.g., for first-order approximations to the advection equation) are 1.085 and 1.843, yielding a convergence factor per single W cycle of 0.27, whereas the single-parameter optimization yields 0.41.

4.5. Numerical experiments. Numerical experiments were carried out with the advection-diffusion equation, discretized by second-order central differences with an added first-order isotropic artificial viscosity term discretized by the five-point star (the importance of the isotropy of the viscosity coefficients is elaborated upon in [6]). The problem solved was

$$-\epsilon \triangle u + \sin(\pi y)\cos(\pi x)\, u_x - \cos(\pi y)\sin(\pi x)\, u_y = 0,$$

over the unit square. The results reported here were obtained with Dirichlet boundary conditions and exact solution zero, in order to allow a very large number of cycles without encountering roundoff errors. The algorithms were tested with smooth nonzero solutions as well, and nearly identical performance was observed until (double-precision) roundoff errors were encountered. Since the physical viscosity coefficient ϵ was taken to be zero, the coefficients of the equation are very small at and near the stagnation point $(\frac{1}{2}, \frac{1}{2})$. When residuals are transferred with some averaging, such as the usual full-weighting that was used in these calculations, the right-hand side on the coarse grid in these regions may be much larger than the coefficients, resulting in reduced performance or even instability. This is best overcome by using averaged coefficients in the calculation of the artificial viscosity. Here we determined the artificial viscosity by adding the absolute value of the coefficient at the point of discretization with weight $\frac{1}{2}$ to those of the four nearest neighbors with weights $\frac{1}{8}$, and multiplying this weighted average by $\frac{h}{2}$ as usual. This averaging introduces only an $O(h^3)$ change in the usual artificial viscosity.

The asymptotic convergence factors of the dynamic residuals with various overweighting factors η are presented in Table 1 along with the predictions of the analyses presented above. The fine mesh-size for the two-level results is 1/64. The multilevel results were obtained with four levels, the finest mesh-size being 1/128, and the analytical prediction refers to a four-level (not infinite-level) cycle, and is calculated from (13). The optimal η is then 1.40—slightly smaller than the infinite-level optimum of $\sqrt{2}$.

The multilevel results were calculated with a W(2,1) cycle (a W cycle with two pre- and one postrelaxation per level), and it was verified that increasing the number of relaxation sweeps per level results in only a negligible improvement in the performance, so that indeed three sweeps per level reduce the high-frequency error components sufficiently, and the convergence rate is determined by the coarse-grid correction.

TABLE 1

Asymptotic residual convergence factors of numerical calculations for the advection-diffusion (AD) *equation and incompressible Navier–Stokes* (INS) *equations are compared with analytical predictions for various single overweighting factors η. Two pre- and one postrelaxation sweeps were employed.*

η	Two-level			Multilevel		
	Analytical prediction	Numerical AD	Numerical INS	Analytical prediction	Numerical AD	Numerical INS
1.00	0.50	0.50	0.49	0.69	0.68	0.66
1.10	0.45	0.45	0.45	0.62	0.61	0.61
1.20	0.40	0.39	0.40	0.55	0.54	0.54
1.30	0.35	0.34	0.33	0.47	0.46	0.45
1.33	0.33	0.33	0.33	0.44	0.44	0.46
1.40	0.40	0.38	0.40	0.40	0.40	0.51
1.50	0.50	0.45	0.51	0.50	0.50	0.63

The experiments were repeated with the incompressible Navier–Stokes (INS) equations in two dimensions, over the unit square with a square of side 0.25 removed from its center. Both the inner and outer squares' sides were aligned with the grid. Dirichlet boundary conditions for the velocities were specified at the inner and outer boundaries. The velocities normal to the boundaries were all made to vanish. The tangential velocities at the inner boundaries were set to zero as well, but the tangential velocities at the outer boundaries were prescribed to be

$$U_{\tan} = \sin \pi s \,,$$

with s varying from 0 to 1 along each side of the outer square, and $U_{\tan}$ driving the flow in the clockwise direction. These conditions and this domain yield a smooth flow with closed streamlines and almost no boundary layers. We chose such a flow because boundary layers constitute a separate problem that needs to be treated by its own specialized methods. Slow convergence due to poor resolution on coarse grids is unrelated to the problem of poor convergence of certain *smooth* components, which is examined here.

The cycle parameters chosen were the same as for the advection-diffusion (AD) equation tests. The discretization used in [5] and [2] was employed with first-order isotropic artificial viscosity, and the Reynolds number solved with was 10^{-6}, so that the physical viscosity was everywhere negligible relative to the artificial viscosity. This is implied by the analysis (and verified by experiments) to be the most difficult case with respect to rate of convergence. Adding physical viscosity, while fixing the overweighting factors, *always* resulted in improved convergence rates.

Now the number of cycles performed was limited by the double-precision roundoff errors. Distributive Gauss–Seidel relaxation with red-black ordering (see [2]) was used throughout. These results are also listed in Table 1 and compared with the analytical prediction for the AD equation.

The numerical results match the analytical prediction very closely. The only exceptions are INS multilevel results with large η's. These, and also the corresponding results in Table 2, indicate greater sensitivity to large overweighting factors in the solution of the system.

Note that *only the momentum equation residuals need to be overweighted.* This is due to the fact that the continuity equation contributes an elliptic component (the Laplacian operator) to the system, for which overweighting is unnecessary. This point has been verified by two-level analysis of the linearized incompressible flow equations.

Experiments were also carried out with the advection-diffusion problem using two different η's. The average two-level convergence factor was 0.34 with the theoretically optimal two-level overweighting factors, and the average multilevel convergence factor was 0.43 when two prerelaxation sweeps and one postrelaxation sweep were employed. The large overweighting factors apparently cause problems. One of these is that overweighting undesirably amplifies high-frequency residuals as well as low-frequency ones, so more smoothing is necessary. Also, since the coarse-grid correction is improved, still more smoothing is required in order to reduce high-frequency errors sufficiently for the smaller convergence factor. Indeed, better performance was obtained when the number of relaxation sweeps per level was increased. With two prerelaxation and two postrelaxation sweeps, the convergence factor attained with the two-level cycle was 0.26 (rather than the theoretical 0.24), and the multilevel factor was 0.37 (rather than the theoretical 0.27). It seems that the assumptions of the analysis are no longer valid, since now intermediate eigenvalues also figure in the convergence rates, and these are neither real nor grid independent.

It is possible to find somewhat smaller overweighting factors that may provide improved results. This requires either a more sophisticated analysis that takes into account high-frequency phenomena, or numerical experimentation in the form of an automatic acceleration process. But in view of the results presented below, this does not seem to be the most profitable course.

5. The defect-correction W (DCW) cycle. The method of defect-correction iterations is a well-known tool for obtaining solutions with high-order accurate operators that are unstable, and that therefore cannot be used directly. This method can be embedded into the multigrid cycle as a means of accelerating convergence.

The basic defect-correction method is as follows. Suppose we wish to obtain an approximate solution to some equation

$$Lu = f \tag{36}$$

with suitable boundary conditions. We would like to use some finite difference operator L_2^h for approximating the differential operator L on a grid with mesh-size h, but cannot use it directly, say, due to problems of instability. Instead, we use another (usually lower-order) operator L_1^h, which is stable, and we hope to approach the L_2^h approximation via the following iterative process:

$$L_1^h u_i^h = f^h + (L_1^h - L_2^h) u_{i-1}^h , \tag{37}$$

where $i \geq 1$ is the iteration number and $u_0^h = 0$. Here again the h superscripts denote functions and operators defined on grid h. If this process converges, it must clearly converge to the solution with the operator L_2^h. In elliptic cases the convergence is usually fast for smooth solution components, in terms of which L_1^h is indeed a good approximation to L_2^h, while the slow convergence of nonsmooth components may actually be an advantage, since for them L_1^h may be better than L_2^h.

A form of this defect-correction process can be used for accelerating the convergence rate when it is slowed down by the poor coarse-grid correction to certain components. We present this technique here as employed in a W cycle, but the generalization to greater cycle indices γ is straightforward.

The two-level *correction scheme* DCW cycle for the solution of the discrete problem

$$L^h u^h = f^h \tag{38}$$

is defined as follows:

1. Start with some initial approximation to u^h on the fine grid h. Smooth the error corresponding to this approximation by some number of relaxation sweeps, obtaining $\tilde{u}^h$.

2. Transfer residuals to the coarse grid H, and solve the first coarse-grid problem

$$L^H v_1^H = r^H \equiv I_h^H \left(f^h - L^h \tilde{u}^h\right) , \tag{39}$$

where the H superscripts denote operators and functions defined on the coarse grid and I_h^H is some transfer operator from grid h to grid H.

3. Calculate and add to the right-hand side a defect-correction term, and solve the second coarse-grid problem:

$$L^H v_2^H = r^H + (L^H - L_h^H) v_1^H , \tag{40}$$

where L_h^H is some approximation to L^h on grid H, which is significantly better than L^H, as explained below.

4. Interpolate and add the correction v_2^H to $\tilde{u}^h$, and smooth this new approximation with some number of relaxation sweeps.

The operator L_h^H used in (40) is some higher-order grid H approximation to the fine-grid operator L^h, which is not used directly, e.g., because it cannot be smoothed efficiently. For example, if L^h and L^H are first-order upstream discretizations of the advection operator on grid h and H, respectively, then they can be viewed as second-order central difference approximations with added artificial viscosity that is proportional to the mesh-size. L_h^H can then be chosen to be a central-difference discretization plus artificial viscosity with coefficients that correspond to grid h, all discretized on grid H. This is then a second-order approximation to L^h. More generally, if L^h is an approximation of order q on grid h to the differential operator L, and, similarly, L^H is a qth-order approximation to L (and L^h) on grid H, then L_h^H can be any operator on grid H which is (at least) a $q+1$st-order approximation to L^h.

The step described in (40) is an attempt to improve the correction yielded by the coarse-grid operator. An alternative view comes to light when $L^H v_1^H$ is subtracted from both sides of (40), yielding

$$L^H (v_2^H - v_1^H) = r^H - L_h^H v_1^H . \tag{41}$$

Viewed thus, the DCW acts as two cycles, except that the visit to the fine grid between the cycles has been skipped, and the process of adding v_1^H to the initial fine-grid approximation and recalculating the residuals is approximated on the coarse grid instead. This viewpoint will later be useful in the analysis.

The multilevel DCW correction cycle is similarly defined, except that the coarse-grid problems are not solved exactly, but rather by a similar DCW cycle (each) on the coarser grid. This is done recursively, and only on the coarsest grid are the problems solved exactly. Note that the cycle index is two (W cycle), which is the reason for the name DCW.

5.1. DCW with FAS. The DCW cycle can of course be implemented with the full approximation scheme (FAS), but it is important to note that the defect correction process must be applied only to the *correction* given by the first "leg" of the W cycle, and not to the full solution. As in the usual FAS algorithm, (39) is now replaced by

$$L^H u_1^H = L^H \hat{I}_h^H \tilde{u}^h + r^H , \tag{42}$$

where $\hat{I}_h^H$ is some transfer operator from grid h to grid H, which need not be the same as I_h^H. Equation (40) is replaced by

$$L^H u_2^H = L^H \hat{I}_h^H \tilde{u}^h + r^H + (L^H - L_h^H)(u_1^H - \hat{I}_h^H \tilde{u}^h) . \tag{43}$$

Finally, $u_2^H - \hat{I}_h^H \tilde{u}^h$ is interpolated and added to $\tilde{u}^h$ as in the usual FAS multigrid cycle. It is easy to verify that for a linear problem this process is equivalent to the correction-scheme two-level DCW cycle.

5.2. Two-level local mode analysis. We analyze the DCW cycle by the same infinite-space analysis employed in the previous sections, again assuming that the factor that determines the rate of convergence is the coarse-grid correction of smooth error components, whereas high-frequency error components are reduced sufficiently by relaxation. Also, intergrid transfers are again neglected, and the FDA is employed.

Suppose the current error in our approximation to the solution of (38) is

$$v^h = e^{i(\omega_1 x + \omega_2 y)} . \tag{44}$$

The corresponding first coarse-grid problem for the correction is then

$$L^H v_1^H = -L^h v^h . \tag{45}$$

Let $C = C(\omega_1, \omega_2) = \hat{L}^h(\omega_1, \omega_2)/\hat{L}^H(\omega_1, \omega_2)$, as in the previous sections. Now

$$v_1^H = -Cv^h ,$$

and the second coarse-grid problem, following (41), is

$$L^H(v_2^H - v_1^H) = -L^h v^h (1 - C) , \tag{46}$$

where no distinction has been made between the actual fine-grid operator L^h and its coarse-grid approximation L_h^H, since the two are equivalent under the FDA. The solution to (46) is

$$v_2^H - v_1^H = -C(1 - C)v^h ,$$

and the error-amplification factor ν after the correction has been added to the fine-grid solution is

$$\nu = \frac{v_{new}^h}{v^h} = \frac{v^h + v_2^H}{v^h} = (1 - C)^2 . \tag{47}$$

This result is quite expected from the point of view of skipping the fine grid. That is, under the present simplifying assumptions the error-amplification factor per cycle is the same as that per two regular cycles as calculated in §2. In particular, the predicted two-level convergence factor per cycle for the first-order discretized advection-diffusion equation is 0.25.

5.3. Multilevel local mode analysis. Following the notation of §3 and the fine-grid skipping viewpoint, we find that, for the multilevel DCW cycle, the one-dimensional map describing the error-amplification factor ν^k (see derivation of (13) with $\gamma = 2$ for the regular W cycle) is given by

$$\nu^k = \left[1 - C(1 - \nu^{k-1})\right]^2 \tag{48}$$

for $k > 0$, with $\nu^0 = 0$. γ is omitted in the notation, since we are only considering a W cycle. Observe now that (48) could also have been obtained from (13) with $\gamma = 2$ by squaring both sides and substituting ν^k and ν^{k-1} for $(\nu_2^k)^2$ and $(\nu_2^{k-1})^2$. So the multilevel DCW cycle is seen to be equivalent to two regular W cycles under the assumptions of this analysis.

Conclusion. The convergence factor of the multilevel DCW cycle for the advection-diffusion equation with first-order discretization ($\xi = 0.5$) tends to one as the number of levels tends to infinity.

5.4. Combining the DCW cycle with OWR. We have seen that the DCW cycle itself does not solve the problem of poor convergence rates when the number of levels is very large, although the rate is squared for a given number of levels. A natural approach is to try to incorporate residual overweighting into the DCW cycle.

5.4.1. η-optimization. The proper way of applying residual overweighting in the DCW cycle is again best seen from the fine-grid skipping point of view. So considered, it is clear that not only do the residuals need to be multiplied by the overweighting factor η, but so does the term $-L_h^H v_1^H$ in (41), which approximates the term that would have been added to the fine-grid residual had the fine grid been visited. By (47), the resulting two-level error-amplification factor with single-parameter overweighting is given by

$$\nu_{tl} = (1 - \eta\xi)^2, \tag{49}$$

where ξ is again the minimal $\hat{L}^h/\hat{L}^H$.

The optimal η is of course the same as that for the regular cycle given in (21) and, by (22), ν_{tl} with the optimal two-level overweighting factor is given by

$$\nu_{tl} = \left(\frac{1-\xi}{1+\xi}\right)^2. \tag{50}$$

This yields a very satisfactory two-level convergence factor per cycle of 0.11 for the first-order discretized advection-diffusion equation.

The optimal multilevel overweighting factor η_{ml} is also the same as for the regular W cycle ($\xi^{-1/2}$), and the multilevel convergence factor, following (27)–(28), is

$$\nu_{ml} = (1 - \xi^{-1/2})^2. \tag{51}$$

For the first-order discretized advection-diffusion equation, $\xi = 0.5$ and $\nu_{ml} = 0.17$.

Multiparameter optimization is also performed as in the case of a regular cycle.

5.5. Defect-correcting for finer levels. In the DCW cycle as presented above, the defect corrections were all used to improve the approximation to the problem on the next-finer grid. But since it is only the *finest*-grid problem whose solution is sought, and not those of the intermediate grids, faster convergence may be obtained by employing operators of a much finer grid in the defect-correction stage. Clearly

this will not affect two-level performance, since then the finest grid is used in the defect correction anyway, but multilevel performance may be improved.

Many schemes are possible, and we consider here the ultimate one of using the operator of the finest grid in all the defect corrections. Under the assumptions of the present analysis, in which the effect of relaxation is disregarded, this scheme is equivalent to employing *double discretization* (see [2]) within a W cycle. (In the double discretization method, *all* the residuals are calculated with the fine-grid operator, but since the initial solution on intermediate levels is zero and the effect of relaxation is disregarded, it does not matter which operator is used on the first leg of the W cycle.) Since now the scheme used on each grid depends on the finest mesh-size, or rather on the ratio of the current mesh-size to the finest one, the multilevel local mode analysis produces a family of recurrence equations rather than just one. Let $k = n - i$ again denote the difference between the finest level n and the solution level i, and let $\nu^{j,k}$ denote the error-amplification factor at level $i + j$ when the finest level is $n = i + k$. Suppose that the error in the approximation to the equation on level $i + j$ is v^j. Then the first leg of the DCW cycle produces a correction v_1^{j-1} of $-C(1 - \nu^{j-1,k})v^j$, as in the regular W cycle, but the second leg now produces an additional correction $v_2^{j-1} - v_1^{j-1}$ of $C(1 - \nu^{j-1,k})[-1 + C^{k-j+1}(1 - \nu^{j-1,k})]v^j$, rather than $C(1 - \nu^{j-1,k})[-1 + C(1 - \nu^{j-1,k})]v^j$ of the usual DCW cycle, since the ratio of the symbol of the operator on level $i + j - 1$ to that of the finest-level operator is C^{k-j+1}. Hence, $\nu^{j,k}$ is given by

$$\nu^{j,k} = 1 - 2C(1 - \nu^{j-1,k}) + C^{k-j+2}(1 - \nu^{j-1,k})^2 , \tag{52}$$

for $j \geq 1$ with $\nu^{0,k} = 0$.

From (52) it is clear that for large values of k some of the $\nu^{j,k}$'s may be quite large in absolute value. But these only express the error-amplification factors for intermediate grids, which are unimportant, because the purpose of the algorithm is to reduce the error on the *finest grid* as efficiently as possible. The only interesting value is therefore $\nu^{k,k}$, which is given by

$$\nu^{k,k} = (1 - C^k)^{2^k} . \tag{53}$$

This result can be proved from (52) by fixing k and performing an induction on j to show that the appropriate binomial expansion is obtained when $j = k$. The proof is omitted due to its length and its irrelevance to the main theme. However, the result is again clear from the fine-grid skipping point of view, although now *all* the levels that are finer than the solution level are skipped. The correction on the solution level is only C^k times the required correction, but 2^k visits are made to this level, in agreement with (53).

PROPOSITION 5. *The maximal* $\nu^{k,k}$ *over all* $\xi \leq C \leq 1$ *and* $k \geq 1$ *tends to* e^{-1} *for* $\xi = 0.5$ *and to* 1 *for* $0 \leq \xi < 0.5$.

Proof. For a fixed k and $0 \leq C \leq 1$, $\nu^{k,k}$ in (53) decreases monotonically as C increases. Therefore, the maximum is obtained when $C = \xi$. Also, for all $k \geq 1$ and $0 < \xi \leq 0.5$,

$$\begin{aligned}\frac{\nu^{k+1,k+1}}{\nu^{k,k}} &= \frac{(1-\xi^{k+1})^{2^{k+1}}}{(1-\xi^k)^{2^k}} \\ &= \left(\frac{(1-\xi\cdot\xi^k)^2}{1-\xi^k}\right)^{2^k} \\ &= \left(\frac{1-2\xi\cdot\xi^k+\xi^{2k+2}}{1-\xi^k}\right)^{2^k} = \left(1+\frac{(1-2\xi)\cdot\xi^k+\xi^{2k+2}}{1-\xi^k}\right)^{2^k} > 1\,.\end{aligned} \tag{54}$$

Hence, the maximal $\nu^{k,k}$ is obtained when $k \to \infty$, and is given by

$$\nu_{ml} = \lim_{k\to\infty}(1-\xi^k)^{2^k} = \lim_{k\to\infty} e^{-(2\xi)^k} = \begin{cases} 1, & 0<\xi<0.5, \\ e^{-1}, & \xi = 0.5. \end{cases} \tag{55}$$

□

Conclusion. The present DCW cycle yields an error-amplification factor of e^{-1} with a multilevel cycle for the first-order discretized advection-diffusion equation.

5.5.1. Combining defect corrections for the finest level with OWR. From (53) it may seem that the method of overweighted residuals cannot be profitably combined with defect corrections for the finest level since, if $C > 1$, $\nu^{k,k}$ diverges as $k \to \infty$. Suppose, however, that the residuals are all transferred with an overweighting factor of η, and the defect-correction term is also multiplied by this factor (rather than by η^{k-j+1}, as is implied by the fine-grid skipping viewpoint). This yields the following recurrence equation for the error-amplification factors $\nu^{j,k}$ (compare with (52)):

$$\nu^{j,k} = 1 - 2\eta C(1-\nu^{j-1,k}) + \eta^2 C^{k-j+2}(1-\nu^{j-1,k})^2\,. \tag{56}$$

PROPOSITION 6. *For all* $0.5 \le C \le 1$ *and all* $k \ge 1$, (56) *with the optimal two-level overweighting factor* $\eta = 4/3$ *satisfies*

$$\nu^{k,k}(C) \le \nu^{1,1}(1) = \tfrac{1}{9}\,. \tag{57}$$

Rather than attempting the difficult task of proving Proposition 6 directly, let us note again that in the present algorithm the solution level can be viewed as solving directly for the correction for the finest grid, rather than for its next-finer grid. But it only yields C^k of the required correction per visit (multiplied by some coefficient that depends on the overweighting factor η). Therefore, $\nu^{k,k}$, which is a polynomial in C with coefficients that depend on η, can also be written as a polynomial in C^k. Let us define accordingly

$$\mu^{k,k}(C^k) \stackrel{\text{def}}{=} \nu^{k,k}(C)\,. \tag{58}$$

$\mu^{k,k}$ also describes the amplification factor of the fine-grid error, but in units of the solution-level correction, rather than the finest-level correction. Therefore, it satisfies a recurrence relationship similar to (56), but with C replaced by 1. In particular, $\mu^{j,k} = \mu^{j,j}$ for all $1 \le j \le k$, yielding the recurrence relationship

$$\mu^{k,k}(C^k) = \left[1-\eta(1-\mu^{k-1,k-1}(C^k))\right]^2 \tag{59}$$

for $k > 1$, with

$$\mu^{1,1}(C^k) = \left(1-\eta C^k\right)^2\,. \tag{60}$$

From this relationship it is fairly straightforward to prove Proposition 6, since the conditions that $\mu^{k,k}$ must satisfy can be traced back to sufficient conditions on $\mu^{1,1}$, which can be shown by induction to be satisfied. We omit the proof since it is irrelevant to the main issues and somewhat lengthy.

Conclusion. Under the present assumptions the optimal two-level convergence factor of 0.11 is attainable with a multilevel cycle.

5.6. Numerical experiments. Numerical experiments were carried out with the same problems as for the method of OWR (§4.5). The fine mesh-size in the two-level experiments was again 1/64, and four levels were employed in the multilevel cycles, the finest mesh-size again being 1/128. The same discretizations were employed.

The operator L_h^H was chosen to be a central-difference approximation (to the advection operator), plus artificial viscosity with coefficients that approximate (to second order) the viscosity coefficients corresponding to grid h.

The results with the regular DCW cycle, as described in §5.4 with various overweighting factors η, appear in Table 2. Two prerelaxation and two postrelaxation sweeps were performed, and the experiments were repeated with three pre- and three postrelaxations (results in parentheses).

TABLE 2

Asymptotic residual convergence factors of numerical calculations for the AD *equation and* INS *equations, solved with a* DCW *cycle, are compared with analytical predictions for various single overweighting factors η. Two (three) pre- and two (three) postrelaxation sweeps were employed.*

η	Two-level			Multilevel		
	Analytical prediction	Numerical AD	Numerical INS	Analytical prediction	Numerical AD	Numerical INS
1.00	0.25	0.25	0.24 (0.24)	0.48	0.48	0.47 (0.46)
1.10	0.20	0.20	0.21 (0.21)	0.38	0.39	0.39 (0.39)
1.20	0.16	0.16	0.21 (0.20)	0.30	0.30	0.32 (0.33)
1.30	0.12	0.11	0.21 (0.19)	0.22	0.22	0.29 (0.27)
1.33	0.11	0.11	0.21 (0.16)	0.20	0.20	0.30 (0.26)
1.40	0.16	0.12	0.24 (0.20)	0.16	0.15	0.33 (0.29)
1.50	0.25	0.20	0.28 (0.25)	0.25	0.20	0.50 (0.50)

The numerical performance of the advection-diffusion solver again matches the prediction well. Performance obtained with large overweighting factors is somewhat better than predicted, especially when more relaxation sweeps are carried out. The reason for this is that, when η is larger than 4/3, the slowest components to converge have the second-finest level as their solution level. For these components, relaxation on the finest grid still has a nonnegligible effect, since their frequencies are fairly high. On the other hand, when η is smaller than 4/3, the slowest-converging components have solution levels that are very coarse, so that even if relaxation on their next-finer grid is very effective, the overall improvement in the convergence rate is very small. Results with the incompressible flow equations were not quite as good as predicted, although a very significant improvement was observed. These results improve further if a small amount of extra artificial viscosity is added, yielding a more effective smoother. When the viscosity coefficients are increased by 40 percent, for example, the asymptotic

residual convergence factors with $\eta = 4/3$ and six relaxation sweeps per level improve to 0.14 (two-level) and 0.24 (multilevel). Of course, there is then a corresponding loss in accuracy, and the cost of performing more sweeps instead must be weighed against the cost of using slightly greater viscosity with a correspondingly finer grid to regain accuracy. Once again, as in the results listed in Table 1, greater overweighting factors than 4/3 yielded poorer results.

The multilevel experiments were repeated with the improvement of correcting the defect for the finest level, as introduced in §5.5. The results are compared with those predicted by numerical analysis of (56) in Table 3. Again, two (three) pre- and two (three) postrelaxations per level were performed. The numerical results match the predictions well for the advection-diffusion equation, although a total of four relaxation sweeps per level (with the present smoother) did not suffice to reduce the high-frequency errors enough. As more and more levels are used, the assumption of the analysis, that the only significant difference between the coarse-grid operators and the fine-grid ones is in the artificial viscosity, becomes poorer and poorer, since the correction level may be much coarser than the finest level. Hence, somewhat better results can be obtained by defect-correcting just a few levels up, and not all the way to the fine-grid operator. An optimal strategy may be worked out experimentally. The INS performance once again lags behind somewhat, but a highly significant improvement is shown, which is again increased by adding some artificial viscosity. When this viscosity is increased by 40 percent, the asymptotic residual convergence factor with $\eta = 4/3$ and six relaxation sweeps per level improves to 0.18. Once again, there is of course a loss of accuracy, which would require a correspondingly finer grid to offset. A slight further increase can be obtained by defect-correcting just a few levels up, rather than for the finest level.

TABLE 3

Asymptotic residual convergence factors of numerical calculations for the AD *equation and* INS *equations, solved with a* DCW *cycle with defect corrections, calculated by the finest-grid operator, are compared with analytical predictions for various single overweighting factors η, which are calculated numerically from* (56). *Two (three) pre- and two (three) postrelaxations were employed.*

η	Analytical prediction	Numerical AD	Numerical INS
1.00	0.37	0.36 (0.36)	0.36 (0.36)
1.10	0.23	0.23 (0.23)	0.28 (0.28)
1.20	0.16	0.16 (0.14)	0.26 (0.23)
1.30	0.12	0.15 (0.10)	0.29 (0.21)
1.33	0.11	0.15 (0.10)	0.31 (0.24)
1.40	0.16	0.20 (0.13)	0.36 (0.34)
1.50	0.25	0.31 (0.21)	0.80 (0.67)

6. Conclusions and remarks. Methods of acceleration of multigrid convergence have been developed, analyzed, and tested, the numerical results mostly matching the predictions of the analyses well. With the optimal method of combining residual overweighting and defect corrections within the W cycle, a multilevel convergence factor of about 0.2 has been obtained for the incompressible flow equations, and 0.1 for the advection-diffusion equations with first-order discretization. We note that large overweighting factors may show poorer behavior than expected for the Navier–Stokes

solver, unless supplemented by extra relaxation sweeps. Some work may be saved by varying the number of sweeps per level.

Although the analyses and experiments were carried out for the case of vanishing diffusion coefficients, the methods apply to finite-viscosity calculations equally well. The optimal parameters are then reduced, and the results improve accordingly, but using the optimal parameters calculated herein will still yield convergence factors that are at least as good as in the vanishing-viscosity case. This has been verified experimentally, the convergence factors actually improving, even with the present optimal overweighting factors, apparently due to improved smoothing. Finally, these methods can also be used with anisotropic viscosity as in upstream differencing.

The methods as presented here are of quite limited value when used *directly* with second-order accurate discretization. Very large overweighting factors need then be applied, which also amplify high-frequency error components. An approach that appears promising is to employ overweighting in conjunction with upstream discretization and downstream ordering of relaxation. Early results with the advection-diffusion equation (with closed streamlines) have been successful, but only very simple examples have so far been attempted, and this approach requires extensive research.

Acknowledgment. We wish to thank Dr. Kees Venner, presently of the Weizmann Institute, for suggesting the particular averaging scheme used in the numerical experiments, and for verifying most of the numerical results reported with independent calculations.

REFERENCES

[1] D. BAI AND A. BRANDT, *Local mesh refinement multilevel techniques*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 109–134.

[2] A. BRANDT, 1984 *Multigrid Guide with Applications to Fluid Dynamics*, Monograph, GMD-Studie 85, GMD-FIT, St. Augustin, West Germany, 1985.

[3] ———, *Multigrid Solvers for Non-Elliptic and Singular-Perturbation Steady-State Problems*, Report, Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel, 1981.

[4] ———, *Rigorous local mode analysis of multigrid*, in Preliminary Proceedings, 4th Copper Mountain Conf. on Multigrid Methods, Chapter 2, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1989.

[5] A. BRANDT AND N. DINAR, *Multigrid solutions to elliptic flow problems*, in Numerical Methods for Partial Differential Equations 53, S. Parter, ed., Academic Press, 1979; ICASE Report 79–15, NASA Langley Research Center, Hampton, VA, 1979.

[6] A. BRANDT AND I. YAVNEH, *Inadequacy of first-order upwind difference schemes for some recirculating flows*, J. Comput. Phys., 93 (1991), pp. 128–143.

[7] ———, *On multigrid solution of high-Reynolds incompressible entering flows*, J. Comput. Phys., 101 (1992), pp. 151–164.

[8] P. M. DE ZEEUW AND E. J. VAN ASSELT, *The convergence rate of multi-level algorithms applied to the convection-diffusion equation*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 492–503.

[9] L. A. HAGEMAN AND D. M. YOUNG, *Applied Iterative Methods*, Academic Press, New York, 1981.

[10] R. MCLACHLAN, *Separated Viscous Flows via Multigrid*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1990.

[11] N. N. YANENKO AND Y. I. SHOKIN, *On the correctness of first differential approximation of difference schemes*, Dokl. Akad. Nauk. SSSR, 182 (1968), pp. 776–778.