Chapter 3

Elements of Multigrid

Through analysis and experimentation, we have examined some of the basic iterative methods. Our discoveries have formed the beginnings of what we might call a spectral (or Fourier mode) picture of relaxation schemes. As we proceed, more essential details of this picture will become clear. So far we have established that many standard iterative methods possess the smoothing property. This property makes these methods very effective at eliminating the high-frequency or oscillatory components of the error, while leaving the low-frequency or smooth components relatively unchanged. The immediate issue is whether these methods can be modified in some way to make them effective on all error components.

One way to improve a relaxation scheme, at least in its early stages, is to use a good initial guess. A well-known technique for obtaining an improved initial guess is to perform some preliminary iterations on a coarse grid. Relaxation on a coarse grid is less expensive because there are fewer unknowns to be updated. Also, because the convergence factor behaves like $1 - O(h^2)$, the coarse grid will have a marginally improved convergence rate. This line of reasoning at least suggests that coarse grids might be worth considering.

With the coarse grid idea in mind, we can think more carefully about its implications. Recall that most basic relaxation schemes suffer in the presence of smooth components of the error. Assume that a particular relaxation scheme has been applied until only smooth error components remain. We now ask what these smooth components look like on a coarser grid. Figure 3.1 shows the answer. A smooth wave with $k = 4$ on a grid Ω^h with $n = 12$ points has been projected directly to the grid Ω^{2h} with $n = 6$ points. On this coarse grid, the original wave still has a wavenumber of $k = 4$. We see that a smooth wave on Ω^h looks *more oscillatory* on Ω^{2h}.

To be more precise, note that the grid points of the coarse grid Ω^{2h} are the even-numbered grid points of the fine grid Ω^h. Consider the kth mode on the fine grid evaluated at the even-numbered grid points. If $1 \le k < \frac{n}{2}$, its components may be written as

$$w^h_{k,2j} = \sin\left(\frac{2jk\pi}{n}\right) = \sin\left(\frac{jk\pi}{n/2}\right) = w^{2h}_{k,j}, \qquad 1 \le k < \frac{n}{2}\,.$$

Notice that superscripts have been used to indicate the grids on which the vectors are defined. From this identity, we see that the kth mode on Ω^h becomes the kth

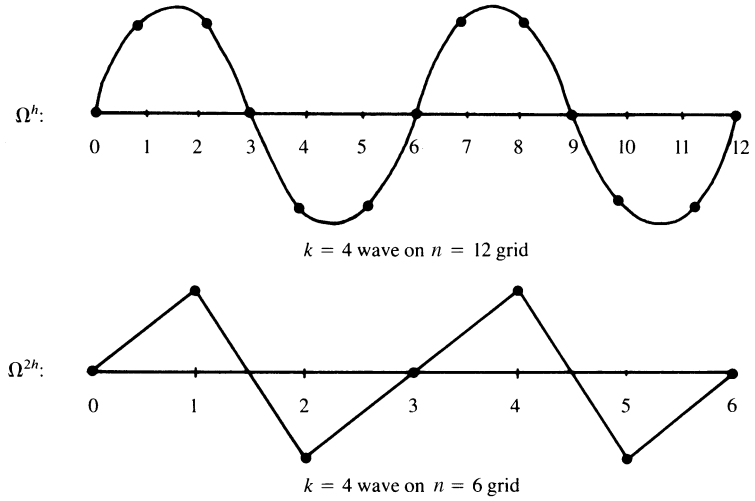

Figure 3.1: *Wave with wavenumber $k = 4$ on Ω^h ($n = 12$ points) projected onto Ω^{2h} ($n = 6$ points). The coarse grid "sees" a wave that is more oscillatory on the coarse grid than on the fine grid.*

mode on Ω^{2h}; this fact is easier to understand by noting that there are half as many modes on Ω^{2h} as there are on Ω^h. The important consequence of this fact is that in passing from the fine grid to the coarse grid, a mode becomes more oscillatory. This is true provided that $1 \leq k < \frac{n}{2}$. It should be verified that the $k = \frac{n}{2}$ mode on Ω^h becomes the zero vector on Ω^{2h}.

As an aside, it is worth mentioning that fine-grid modes with $k > \frac{n}{2}$ undergo a more curious transformation. Through the phenomenon of *aliasing* mentioned earlier, the kth mode on Ω^h becomes the $(n-k)$th mode on Ω^{2h} when $k > \frac{n}{2}$ (Exercise 1). In other words, the oscillatory modes of Ω^h are misrepresented as relatively smooth modes on Ω^{2h}.

The important point is that smooth modes on a fine grid look less smooth on a coarse grid. This suggests that when relaxation begins to stall, signaling the predominance of smooth error modes, it is advisable to move to a coarser grid; there, the smooth error modes appear more oscillatory and relaxation will be more effective. The question is: how do we move to a coarser grid and relax on the more oscillatory error modes?

It is at this point that multigrid begins to come together like a jigsaw puzzle. We must keep all of the related facts in mind. Recall that we have an equation for the error itself, namely, the residual equation. If $\mathbf{v}$ is an approximation to the exact solution $\mathbf{u}$, then the error $\mathbf{e} = \mathbf{u} - \mathbf{v}$ satisfies

$$A\mathbf{e} = \mathbf{r} = \mathbf{f} - A\mathbf{v},$$

which says that we can relax directly on the error by using the residual equation. There is another argument that justifies the use of the residual equation:

> Relaxation on the original equation $A\mathbf{u} = \mathbf{f}$ with an arbitrary initial guess $\mathbf{v}$ is equivalent to relaxing on the residual equation $A\mathbf{e} = \mathbf{r}$ with the specific initial guess $\mathbf{e} = \mathbf{0}$.

This intimate connection between the original and the residual equations further motivates the use of the residual equation (Exercise 2).

We must now gather these loosely connected ideas. We know that many relaxation schemes possess the smoothing property. This leads us to consider using coarser grids during the computation to focus the relaxation on the oscillatory components of the error. In addition, there seems to be good reason to involve the residual equation in the picture. We now try to give these ideas a little more definition by proposing two strategies.

We begin by proposing a strategy that uses coarse grids to obtain better initial guesses.

- Relax on $A\mathbf{u} = \mathbf{f}$ on a very coarse grid to obtain an initial guess for the next finer grid.

  .
  .
  .

- Relax on $A\mathbf{u} = \mathbf{f}$ on Ω^{4h} to obtain an initial guess for Ω^{2h}.
- Relax on $A\mathbf{u} = \mathbf{f}$ on Ω^{2h} to obtain an initial guess for Ω^{h}.
- Relax on $A\mathbf{u} = \mathbf{f}$ on Ω^{h} to obtain a final approximation to the solution.

This idea of using coarser grids to generate improved initial guesses is the basis of a strategy called *nested iteration*. Although the approach is attractive, it also leaves some questions. For instance, what does it mean to relax on $A\mathbf{u} = \mathbf{f}$ on Ω^{2h}? We must somehow define the original problem on the coarser grids. Also, what happens if, having once reached the fine grid, there are still smooth components in the error? We may have obtained some improvement by using the coarse grids, but the final iteration will stall if smooth components still remain. We return to these questions and find answers that will allow us to use nested iteration in a very powerful way.

A second strategy incorporates the idea of using the residual equation to relax on the error. It can be represented by the following procedure:

- Relax on $A\mathbf{u} = \mathbf{f}$ on Ω^{h} to obtain an approximation $\mathbf{v}^h$.
- Compute the residual $\mathbf{r} = \mathbf{f} - A\mathbf{v}^h$.

  Relax on the residual equation $A\mathbf{e} = \mathbf{r}$ on Ω^{2h} to obtain an approximation to the error $\mathbf{e}^{2h}$.

- Correct the approximation obtained on Ω^{h} with the error estimate obtained on Ω^{2h} : $\mathbf{v}^h \leftarrow \mathbf{v}^h + \mathbf{e}^{2h}$.

This procedure is the basis of what is called *the correction scheme*. Having relaxed on the fine grid until convergence deteriorates, we relax on the *residual* equation on a coarser grid to obtain an approximation to the error itself. We then return to the fine grid to correct the approximation first obtained there.

There is a rationale for using this correction strategy, but it also leaves some questions to be answered. For instance, what does it mean to relax on $A\mathbf{e} = \mathbf{r}$ on Ω^{2h}? To answer this question, we first need to know how to compute the residual

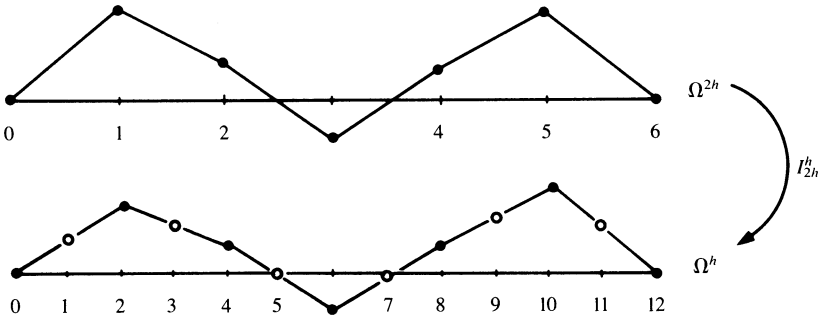

Figure 3.2: *Interpolation of a vector on coarse grid Ω^{2h} to fine grid Ω^h.*

on Ω^h and transfer it to Ω^{2h}. We also need to know how to relax on Ω^{2h} and what initial guess should be used. Moreover, how do we transfer the error estimate from Ω^{2h} back to Ω^h? These questions suggest that we need mechanisms for transferring information between the grids. We now turn to this important consideration.

In our discussion of intergrid transfers, we consider only the case in which the coarse grid has twice the grid spacing of the next finest grid. This is a nearly universal practice, because there is usually no advantage in using grid spacings with ratios other than 2. Think for a moment about the step in the correction scheme that requires transferring the error approximation $\mathbf{e}^{2h}$ from the coarse grid Ω^{2h} to the fine grid Ω^h. This is a common procedure in numerical analysis and is generally called *interpolation* or *prolongation*. Many interpolation methods could be used. Fortunately, for most multigrid purposes, the simplest of these is quite effective. For this reason, we consider only linear interpolation.

The linear interpolation operator will be denoted I_{2h}^h. It takes coarse-grid vectors and produces fine-grid vectors according to the rule $I_{2h}^h \mathbf{v}^{2h} = \mathbf{v}^h$, where

$$\begin{aligned} v_{2j}^h &= v_j^{2h}, \\ v_{2j+1}^h &= \frac{1}{2}\left(v_j^{2h} + v_{j+1}^{2h}\right), \quad 0 \le j \le \frac{n}{2} - 1. \end{aligned}$$

Figure 3.2 shows graphically the action of I_{2h}^h. At even-numbered fine-grid points, the values of the vector are transferred directly from Ω^{2h} to Ω^h. At odd-numbered fine-grid points, the value of $\mathbf{v}^h$ is the average of the adjacent coarse-grid values.

In anticipation of discussions to come, we note that I_{2h}^h is a linear operator from $\mathbf{R}^{\frac{n}{2}-1}$ to $\mathbf{R}^{n-1}$. It has full rank and the trivial null space, $\mathcal{N} = \{0\}$. For the case $n = 8$, this operator has the form

$$I_{2h}^h \mathbf{v}^{2h} = \frac{1}{2}\begin{bmatrix} 1 & & \\ 2 & & \\ 1 & 1 & \\ & 2 & \\ & 1 & 1 \\ & & 2 \\ & & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_{2h} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}_h = \mathbf{v}^h.$$

How well does this interpolation process work? First assume that the "real" error (which is not known exactly) *is* a smooth vector on the fine grid. Assume

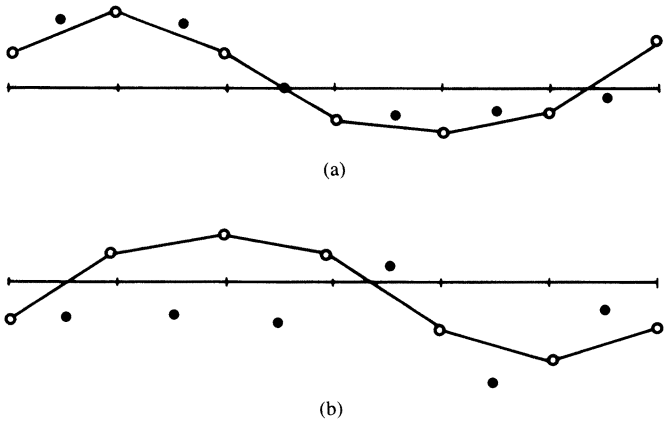

Figure 3.3: (a) *If the exact error on* Ω^h *(indicated by* $\circ$ *and* $\bullet$*) is smooth, an interpolant of the coarse-grid error* $\mathbf{e}^{2h}$ *(solid line connecting* $\circ$ *points) should give a good representation of the exact error.* (b) *If the exact error on* Ω^h *(indicated by* $\circ$ *and* $\bullet$*) is oscillatory, an interpolant of the coarse-grid error* $\mathbf{e}^{2h}$ *(solid line connecting* $\circ$ *points) may give a poor representation of the exact error.*

also that a coarse-grid approximation to the error has been determined on Ω^{2h} and that this approximation is exact at the coarse-grid points. When this coarse-grid approximation is interpolated to the fine grid, the interpolant is also smooth. Therefore, we expect a relatively good approximation to the fine-grid error, as shown in Fig. 3.3(a). By contrast, if the "real" error is oscillatory, even a very good coarse-grid approximation may produce an interpolant that is not very accurate. This situation is shown in Fig. 3.3(b).

Thus, interpolation is most effective when the error is smooth. Because interpolation is necessary for both nested iteration and the correction scheme, we may conclude that these two processes are most effective when the error is smooth. As we will see shortly, these processes provide a fortunate complement to relaxation, which is most effective when the error is oscillatory.

For two-dimensional problems, the interpolation operator may be defined in a similar way. If we let $I_{2h}^h \mathbf{v}^{2h} = \mathbf{v}^h$, then the components of $\mathbf{v}^h$ are given by

$$\begin{aligned}
v_{2i,2j}^h &= v_{ij}^{2h}, \\
v_{2i+1,2j}^h &= \frac{1}{2}\left(v_{ij}^{2h} + v_{i+1,j}^{2h}\right), \\
v_{2i,2j+1}^h &= \frac{1}{2}\left(v_{ij}^{2h} + v_{i,j+1}^{2h}\right), \\
v_{2i+1,2j+1}^h &= \frac{1}{4}\left(v_{ij}^{2h} + v_{i+1,j}^{2h} + v_{i,j+1}^{2h} + v_{i+1,j+1}^{2h}\right), \quad 0 \le i,j \le \frac{n}{2} - 1.
\end{aligned}$$

The second class of intergrid transfer operations involves moving vectors from a fine grid to a coarse grid. They are generally called *restriction* operators and are denoted by I_h^{2h}. The most obvious restriction operator is *injection*. It is defined by $I_h^{2h}\mathbf{v}^h = \mathbf{v}^{2h}$, where

$$v_j^{2h} = v_{2j}^h.$$

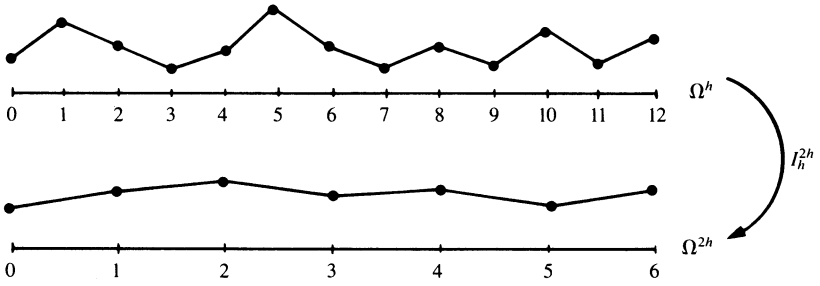

Figure 3.4: *Restriction by full weighting of a fine-grid vector to the coarse grid.*

In other words, with injection, the coarse-grid vector simply takes its value directly from the corresponding fine-grid point.

An alternate restriction operator, called *full weighting*, is defined by $I_h^{2h}\mathbf{v}^h = \mathbf{v}^{2h}$, where

$$v_j^{2h} = \frac{1}{4}\big(v_{2j-1}^h + 2v_{2j}^h + v_{2j+1}^h\big), \qquad 1 \le j \le \frac{n}{2} - 1.$$

As Fig. 3.4 shows, the values of the coarse-grid vector are weighted averages of values at neighboring fine-grid points.

In the discussion that follows, we use full weighting as a restriction operator. However, in some instances, injection may be the better choice. The issue of intergrid transfers, which is an important part of multigrid theory, is discussed at some length in Brandt's guide to multigrid [4].

The full weighting operator is a linear operator from $\mathbf{R}^{n-1}$ to $\mathbf{R}^{\frac{n}{2}-1}$. It has a rank of $\frac{n}{2} - 1$ (Exercise 4) and a null space of dimension $\frac{n}{2}$ (Exercise 5). For the case $n = 8$, the full weighting operator has the form

$$I_h^{2h}\mathbf{v}^h = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 & & & & & \\ & & 1 & 2 & 1 & & & \\ & & & & 1 & 2 & 1 & \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}_h = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_{2h} = \mathbf{v}^{2h}.$$

One reason for our choice of full weighting as a restriction operator is the important fact (Exercise 6) that

$$I_{2h}^h = c(I_h^{2h})^T, \qquad c \in \mathbf{R}.$$

The fact that the interpolation operator and the full weighting operator are transposes of each other up to a constant is called a *variational property* and will soon be of importance.

For the sake of completeness, we give the full weighting operator in two dimensions. It is just an averaging of the fine-grid nearest neighbors. Letting $I_h^{2h}\mathbf{v}^h = \mathbf{v}^{2h}$, we have that

$$\begin{aligned} v_{ij}^{2h} \;=\; \frac{1}{16}\big[&v_{2i-1,2j-1}^h + v_{2i-1,2j+1}^h + v_{2i+1,2j-1}^h + v_{2i+1,2j+1}^h \\ &+ 2\big(v_{2i,2j-1}^h + v_{2i,2j+1}^h + v_{2i-1,2j}^h + v_{2i+1,2j}^h\big) \\ &+ 4v_{2i,2j}^h\big], \quad 1 \le i, j \le \frac{n}{2} - 1. \end{aligned}$$

We now have a well-defined way to transfer vectors between fine and coarse grids. Therefore, we can return to the correction scheme and make it precise. To do this, we define the following two-grid correction scheme.

Two-Grid Correction Scheme

$$\mathbf{v}^h \leftarrow MG(\mathbf{v}^h, \mathbf{f}^h).$$

- Relax ν_1 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ on Ω^h with initial guess $\mathbf{v}^h$.
- Compute the fine-grid residual $\mathbf{r}^h = \mathbf{f}^h - A^h\mathbf{v}^h$ and restrict it to the coarse grid by $\mathbf{r}^{2h} = I_h^{2h} r^h$.
- Solve $A^{2h}\mathbf{e}^{2h} = \mathbf{r}^{2h}$ on Ω^{2h}.
- Interpolate the coarse-grid error to the fine grid by $\mathbf{e}^h = I_{2h}^h \mathbf{e}^{2h}$ and correct the fine-grid approximation by $\mathbf{v}^h \leftarrow \mathbf{v}^h + \mathbf{e}^h$.
- Relax ν_2 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ on Ω^h with initial guess $\mathbf{v}^h$.

This procedure is simply the original correction scheme, now refined by the use of the intergrid transfer operators. We relax on the fine grid until it ceases to be worthwhile; in practice, ν_1 is often 1, 2, or 3. The residual of the current approximation is computed on Ω^h and then transferred by a restriction operator to the coarse grid. As it stands, the procedure calls for the exact solution of the residual equation on Ω^{2h}, which may not be possible. However, if the coarse-grid error can at least be approximated, it is then interpolated up to the fine grid, where it is used to correct the fine-grid approximation. This is followed by ν_2 additional fine-grid relaxation sweeps.

Several comments are in order. First, notice that the superscripts h or $2h$ are essential to indicate the grid on which a particular vector or matrix is defined. Second, all of the quantities in the above procedure are well defined except for A^{2h}. For the moment, we take A^{2h} simply to be the result of discretizing the problem on Ω^{2h}. Finally, the integers ν_1 and ν_2 are parameters in the scheme that control the number of relaxation sweeps before and after visiting the coarse grid. They are usually fixed at the start, based on either theoretical considerations or on past experimental results.

It is important to appreciate the complementarity at work in the process. Relaxation on the fine grid eliminates the oscillatory components of the error, leaving a relatively smooth error. Assuming the residual equation can be solved accurately on Ω^{2h}, it is still important to transfer the error accurately back to the fine grid. Because the error is smooth, interpolation should work very well and the correction of the fine-grid solution should be effective.

Numerical example. A numerical example will be helpful. Consider the weighted Jacobi method with $\omega = \frac{2}{3}$ applied to the one-dimensional model problem $A\mathbf{u} = \mathbf{0}$ on a grid with $n = 64$ points. We use an initial guess,

$$v_j^h = \frac{1}{2}\left[\sin\left(\frac{16j\pi}{n}\right) + \sin\left(\frac{40j\pi}{n}\right)\right],$$

consisting of the $k = 16$ and $k = 40$ modes. The following two-grid correction scheme is used:

- Relax three times on $A^h\mathbf{u}^h = \mathbf{0}$ on Ω^h with initial guess $\mathbf{v}^h$.
- Compute $\mathbf{r}^{2h} = I_h^{2h}\mathbf{r}^h$.
- Relax three times on $A^{2h}\mathbf{e}^{2h} = \mathbf{r}^{2h}$ on Ω^{2h} with initial guess $\mathbf{e}^{2h} = \mathbf{0}$.
- Correct the fine-grid approximation: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{e}^{2h}$.
- Relax three times on $A^h\mathbf{u}^h = \mathbf{0}$ on Ω^h with initial guess $\mathbf{v}^h$.
- Compute $\mathbf{r}^{2h} = I_h^{2h}\mathbf{r}^h$.
- Relax three times on $A^{2h}\mathbf{e}^{2h} = \mathbf{r}^{2h}$ on Ω^{2h} with initial guess $\mathbf{e}^{2h} = \mathbf{0}$.
- Correct the fine-grid approximation: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{e}^{2h}$.

The results of this calculation are given in Fig. 3.5. The initial guess with its two modes is shown in the top left figure. In the top right, the approximation $\mathbf{v}^h$ after one relaxation sweep is superimposed on the initial guess. Much of the oscillatory component of the initial guess has already been removed, and the 2-norm of the error has been diminished to 57% of the norm of the initial error. The middle left plot shows the approximation after three relaxation sweeps on the fine grid, again superimposed on the initial guess. The solution (in this case, the error) has become smoother and its norm is now 36% of the initial error norm. Further relaxations on the fine grid would provide only a slow improvement at this point. This signals that it is time to move to the coarse grid.

The middle right plot shows the fine-grid error after one relaxation sweep on the coarse-grid residual equation, superimposed on the initial guess. Clearly, we have achieved another reduction in the error by moving to the coarse grid; the norm of the error is now 26% of the initial error norm. This improvement occurs because the smooth error components, inherited from the fine grid, appear oscillatory on the coarse grid and are quickly removed. The error after three coarse-grid relaxation sweeps is shown in the bottom left figure. The norm of the error is now about 8% of its initial value.

The coarse-grid approximation to the error is now used to correct the fine-grid approximation. After three additional fine-grid relaxations, the 2-norm of the error is reduced to about 3% of the initial error norm. This result is plotted in the bottom right figure. The residual is once again transferred to the coarse grid and three coarse-grid relaxations follow. At this point, the 2-norm of the error is about 1% of its original value. This experiment demonstrates that relaxation, when done on two grids and applied to both the original and the residual equation, can be very powerful. ◇◇

The two-grid correction scheme, as outlined above, leaves one looming procedural question: what is the best way to solve the coarse-grid problem $A^{2h}\mathbf{e}^{2h} = \mathbf{r}^{2h}$? The answer may be apparent, particularly to those who think recursively. The coarse-grid problem is not much different from the original problem. Therefore, we can apply the two-grid correction scheme to the residual equation on Ω^{2h}, which means relaxing there and then moving to Ω^{4h} for the correction step. We can repeat this process on successively coarser grids until a direct solution of the residual equation is possible.

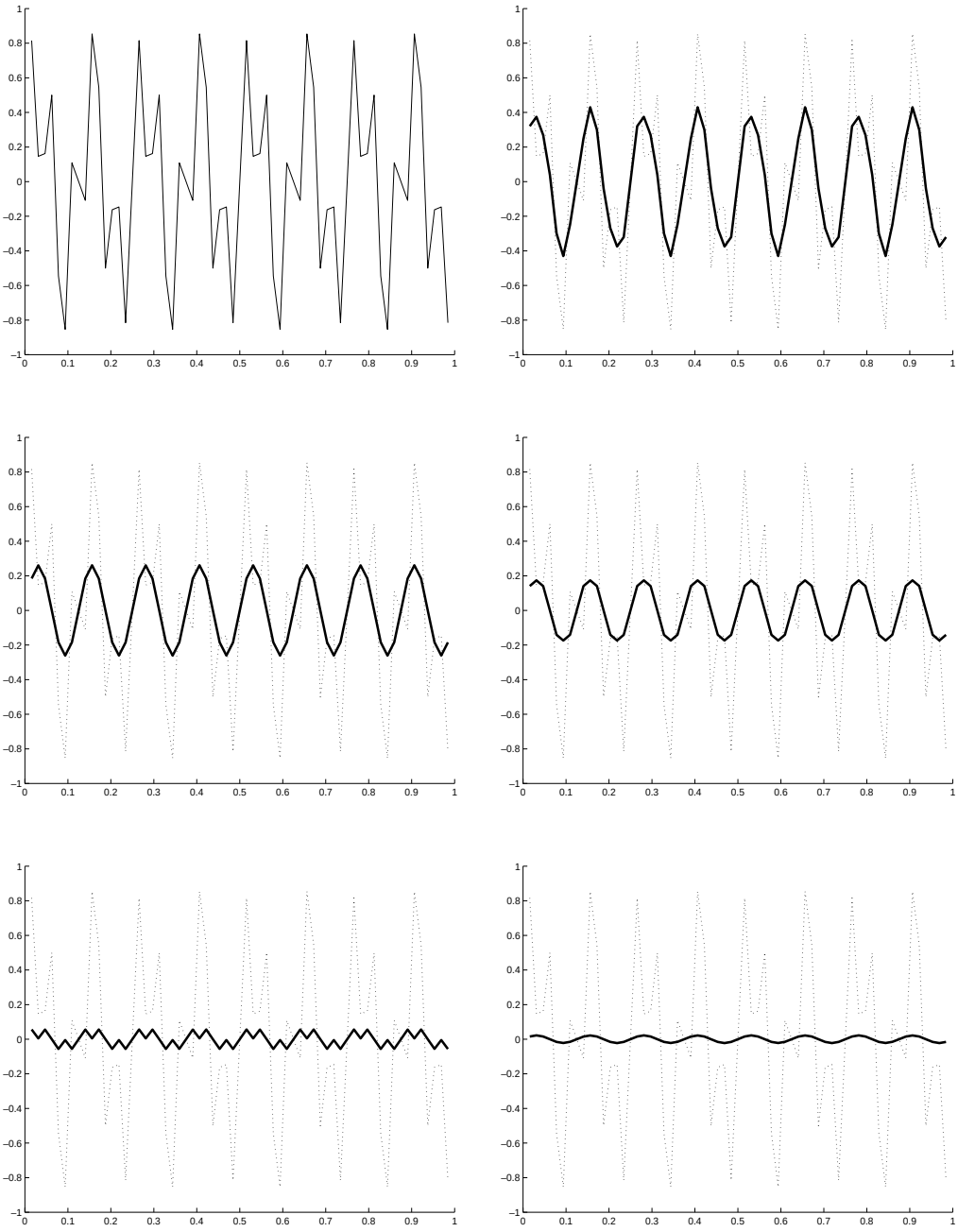

Figure 3.5: *Coarse-grid correction for* $-u'' = 0$ *on a grid with* $n = 64$. Top left: *The initial guess,* $(\mathbf{w}_{16} + \mathbf{w}_{40})/2$. Top right: *The error after one sweep of weighted Jacobi.* Middle left: *The error after three sweeps of weighted Jacobi.* Middle right: *The fine-grid error after one sweep of weighted Jacobi on the coarse-grid problem.* Bottom left: *The fine-grid error after three sweeps of weighted Jacobi on the coarse-grid problem.* Bottom right: *The fine-grid error after the coarse-grid correction is followed by three weighted Jacobi sweeps on the fine grid.*

To facilitate the description of this procedure, some economy of notation is desirable. The same notation is used for the computer implementation of the resulting algorithm. We call the right-side vector of the residual equation $\mathbf{f}^{2h}$, rather than $\mathbf{r}^{2h}$, because it is just another right-side vector. Instead of calling the solution of the residual equation $\mathbf{e}^{2h}$, we use $\mathbf{u}^{2h}$ because it is just a solution vector. We can then use $\mathbf{v}^{2h}$ to denote approximations to $\mathbf{u}^{2h}$. These changes simplify the notation, but it is still important to remember the meaning of these variables.

One more point needs to be addressed: what initial guess do we use for $\mathbf{v}^{2h}$ on the first visit to Ω^{2h}? Because there is presumably no information available about the solution, $\mathbf{u}^{2h}$, we simply choose $\mathbf{v}^{2h} = \mathbf{0}$. Here then is the two-grid correction scheme, now imbedded within itself. We assume that there are $l > 1$ grids with grid spacings $h, 2h, 4h, \ldots, Lh = 2^{l-1}h$.

V-Cycle Scheme

$$\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h)$$

- Relax on $A^h\mathbf{u}^h = \mathbf{f}^h$ ν_1 times with initial guess $\mathbf{v}^h$.
- Compute $\mathbf{f}^{2h} = I_h^{2h}\mathbf{r}^h$.
  - Relax on $A^{2h}\mathbf{u}^{2h} = \mathbf{f}^{2h}$ ν_1 times with initial guess $\mathbf{v}^{2h} = \mathbf{0}$.
  - Compute $\mathbf{f}^{4h} = I_{2h}^{4h}\mathbf{r}^{2h}$.
    - Relax on $A^{4h}\mathbf{u}^{4h} = \mathbf{f}^{4h}$ ν_1 times with initial guess $\mathbf{v}^{4h} = \mathbf{0}$.
    - Compute $\mathbf{f}^{8h} = I_{4h}^{8h}\mathbf{r}^{4h}$.

      .
      .
      .

      - Solve $A^{Lh}\mathbf{u}^{Lh} = \mathbf{f}^{Lh}$.

      .
      .
      .

    - Correct $\mathbf{v}^{4h} \leftarrow \mathbf{v}^{4h} + I_{8h}^{4h}\mathbf{v}^{8h}$.
    - Relax on $A^{4h}\mathbf{u}^{4h} = \mathbf{f}^{4h}$ ν_2 times with initial guess $\mathbf{v}^{4h}$.
  - Correct $\mathbf{v}^{2h} \leftarrow \mathbf{v}^{2h} + I_{4h}^{2h}\mathbf{v}^{4h}$.
  - Relax on $A^{2h}\mathbf{u}^{2h} = \mathbf{f}^{2h}$ ν_2 times with initial guess $\mathbf{v}^{2h}$.
- Correct $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{v}^{2h}$.
- Relax on $A^h\mathbf{u}^h = \mathbf{f}^h$ ν_2 times with initial guess $\mathbf{v}^h$.

The algorithm telescopes down to the coarsest grid, which can consist of one or a few interior grid points, then works its way back to the finest grid. Figure 3.6(a) shows the schedule for the grids in the order in which they are visited. Because of the pattern in this diagram, this algorithm is called the *V-cycle*. It is our first true multigrid method.

Not surprisingly, the V-cycle has a compact recursive definition, which is given as follows.

V-Cycle Scheme (Recursive Definition)

$$\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h).$$

1. Relax ν_1 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ with a given initial guess $\mathbf{v}^h$.

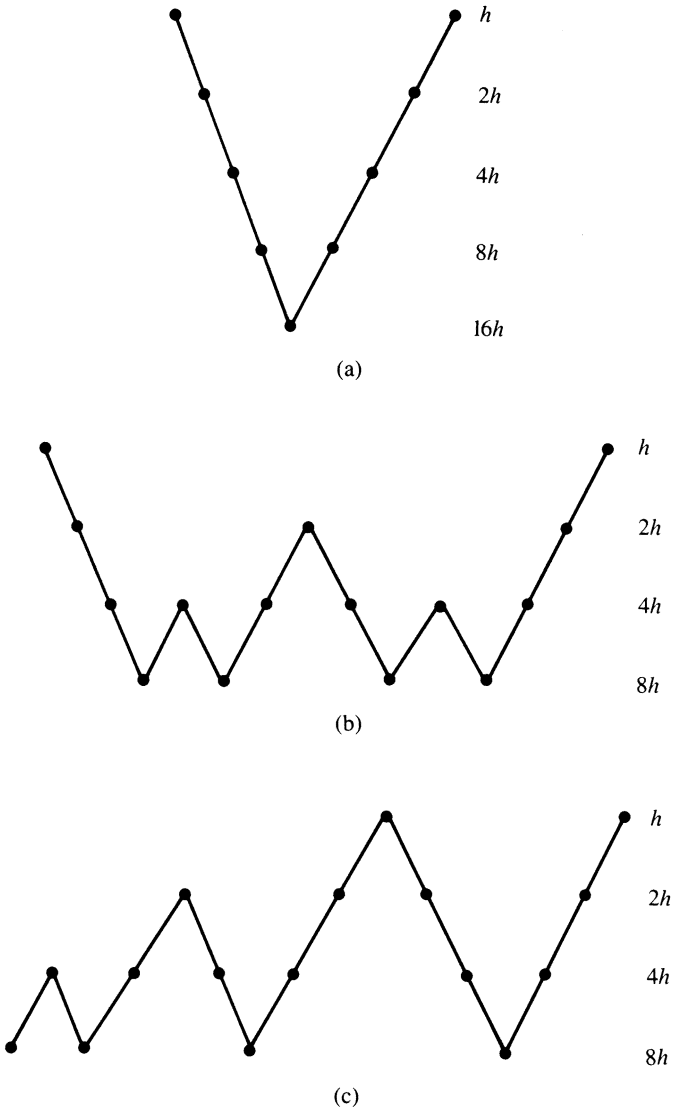

Figure 3.6: *Schedule of grids for* (a) *V-cycle,* (b) *W-cycle, and* (c) *FMG scheme, all on four levels.*

2. If Ω^h = coarsest grid, then go to step 4.
Else

$$\mathbf{f}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h),$$
$$\mathbf{v}^{2h} \leftarrow \mathbf{0},$$
$$\mathbf{v}^{2h} \leftarrow V^{2h}(\mathbf{v}^{2h}, \mathbf{f}^{2h}).$$

3. Correct $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{v}^{2h}$.

4. Relax ν_2 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ with initial guess $\mathbf{v}^h$.

The V-cycle is just one of a family of multigrid cycling schemes. The entire family is called the *μ-cycle* method and is defined recursively by the following.

μ-Cycle Scheme

$$\mathbf{v}^h \leftarrow M\mu^h(\mathbf{v}^h, \mathbf{f}^h).$$

1. Relax ν_1 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ with a given initial guess $\mathbf{v}^h$.

2. If Ω^h = coarsest grid, then go to step 4.

Else

$$\mathbf{f}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h),$$
$$\mathbf{v}^{2h} \leftarrow \mathbf{0},$$
$$\mathbf{v}^{2h} \leftarrow M\mu^{2h}(\mathbf{v}^{2h}, \mathbf{f}^{2h}) \ \mu \text{ times.}$$

3. Correct $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{v}^{2h}$.

4. Relax ν_2 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ with initial guess $\mathbf{v}^h$.

In practice, only $\mu = 1$ (which gives the V-cycle) and $\mu = 2$ are used. Figure 3.6(b) shows the schedule of grids for $\mu = 2$ and the resulting *W-cycle*. We refer to a V-cycle with ν_1 relaxation sweeps before the correction step and ν_2 relaxation sweeps after the correction step as a V(ν_1, ν_2)-cycle, with a similar notation for W-cycles.

We originally stated that two ideas would lead to multigrid. So far we have developed only the correction scheme. The nested iteration idea has yet to be explored. Recall that nested iteration uses coarse grids to obtain improved initial guesses for fine-grid problems. In looking at the V-cycle, we might ask how to obtain an informed initial guess for the first fine-grid relaxation. Nested iteration would suggest solving a problem on Ω^{2h}. But how can we obtain a good initial guess for the Ω^{2h} problem? Nested iteration sends us to Ω^{4h}. Clearly, we are on another recursive path that leads to the coarsest grid.

The algorithm that joins nested iteration with the V-cycle is called the *full multigrid V-cycle (FMG)* . Given first in explicit terms, it appears as follows.

Full Multigrid V-Cycle

$$\mathbf{v}^h \leftarrow FMG^h(\mathbf{f}^h).$$

Initialize $f^{2h} \leftarrow I_h^{2h} f^h$, $f^{4h} \leftarrow I_{2h}^{4h} f^{2h}, \ldots$.

- Solve or relax on coarsest grid.

.
.
.

- $\mathbf{v}^{4h} \leftarrow I_{8h}^{4h}\mathbf{v}^{8h}$.
- $\mathbf{v}^{4h} \leftarrow V^{4h}(\mathbf{v}^{4h}, \mathbf{f}^{4h})\ \nu_0$ times.

- $\mathbf{v}^{2h} \leftarrow I_{4h}^{2h}\mathbf{v}^{4h}$.
- $\mathbf{v}^{2h} \leftarrow V^{2h}(\mathbf{v}^{2h}, \mathbf{f}^{2h})\ \nu_0$ times.

- $\mathbf{v}^h \leftarrow I_{2h}^h\mathbf{v}^{2h}$.
- $\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h), \nu_0$ times.

We initialize the coarse-grid right sides by transferring $\mathbf{f}^h$ from the fine grid. Another option is to use the original right-side function f. The cycling parameter, ν_0, sets the number of V-cycles done at each level. It is generally determined by a previous numerical experiment; $\nu_0 = 1$ is the most common choice. Expressed recursively, the algorithm has the following compact form.

Full Multigrid V-Cycle (Recursive Form)

$$\mathbf{v}^h \leftarrow FMG^h(\mathbf{f}^h).$$

1. If Ω^h = coarsest grid, set $\mathbf{v}^h \leftarrow \mathbf{0}$ and go to step 3.
Else
$$\mathbf{f}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h),$$
$$\mathbf{v}^{2h} \leftarrow FMG^{2h}(\mathbf{f}^{2h}).$$

2. Correct $\mathbf{v}^h \leftarrow I_{2h}^h \mathbf{v}^{2h}$.

3. $\mathbf{v}^h \leftarrow V^h(\mathbf{v}^h, \mathbf{f}^h)$ ν_0 times.

Figure 3.6(c) shows the schedule of grids for FMG with $\nu_0 = 1$. Each V-cycle is preceded by a coarse-grid V-cycle designed to provide the best initial guess possible. As we will see, the extra work done in these preliminary V-cycles is not only inexpensive (Exercise 8), but easily pays for itself.

Full multigrid is the complete knot into which the many threads of the preceding chapters are tied. It is a remarkable synthesis of ideas and techniques that individually have been well known and used for a long time. Taken alone, many of these ideas have serious defects. Full multigrid is a technique for integrating them so that they can work together in a way that overcomes these limitations. The result is a very powerful algorithm.

Exercises

1. **Aliasing.** Show that the kth mode on a grid Ω^h with $n-1$ interior points appears as the $(n-k)$th mode on Ω^{2h} when $\frac{n}{2} < k < n$.

2. **An important equivalence.** Consider a stationary, linear method of the form $\mathbf{v} \leftarrow \mathbf{v} + B^{-1}(\mathbf{f} - A\mathbf{v})$ applied to the problem $A\mathbf{u} = \mathbf{f}$. Use the following steps to show that relaxation on $A\mathbf{u} = \mathbf{f}$ with an arbitrary initial guess is equivalent to relaxation on $A\mathbf{e} = \mathbf{r}$ with the zero initial guess:

(a) First consider the problem $A\mathbf{u} = \mathbf{f}$ with an arbitrary initial guess $\mathbf{v} = \mathbf{v_0}$. What are the error and residual associated with $\mathbf{v_0}$?

(b) Now consider the associated residual equation $A\mathbf{e} = \mathbf{r_0} = \mathbf{f} - A\mathbf{v_0}$. What are the error and residual in the initial guess $\mathbf{e_0} = \mathbf{0}$?

(c) Conclude that the problems in (a) and (b) are equivalent.

3. **Properties of interpolation.** Show that I_{2h}^h based upon linear interpolation is a linear operator with full rank in one and two dimensions.

4. **Properties of restriction.** What is the rank of I_h^{2h} based on (a) full weighting and (b) injection in one and two dimensions?

5. **Null space of full weighting.** Show that the null space of the full weighting operator, $N(I_h^{2h})$, has a basis consisting of vectors of the form

$$(0, 0, \ldots, -1, 2, -1, \ldots, 0, 0)^T.$$

By counting these vectors, show that the dimension of $N(I_h^{2h})$ is $\frac{n}{2}$.

6. **Variational property.**

(a) Let I_{2h}^h and I_h^{2h} be defined as in the text. Show that linear interpolation and full weighting satisfy the variational property $I_{2h}^h = c(I_h^{2h})^T$ by computing $c \in \mathbf{R}$ for both one and two dimensions.

(b) The choice of $c \neq 1$ found in part (a) is used because full weighting essentially preserves constants. Show that, except at the boundary, $I_h^{2h}(\mathbf{1}^h) = \mathbf{1}^{2h}$ (where $\mathbf{1}^h$ and $\mathbf{1}^{2h}$ are the vectors with entries 1 on their respective grids).

7. **Properties of red-black Gauss–Seidel.** Suppose red-black Gauss–Seidel is used with the V-cycle scheme for the one-dimensional model problem.

(a) Does it matter whether the odd unknowns or even unknowns are updated first? Explain.

(b) Show that one sweep of red-black Gauss–Seidel on Ω^h leaves the error $\mathbf{e}^h$ in the range of interpolation I_{2h}^h.

(c) Demonstrate that one V-cycle based on red-black Gauss–Seidel and full weighting is a direct (exact) solver for the one-dimensional model problem.

8. **FMG cost.** The difference in cost between FMG and a single V-cycle is the cost of all but the last V-cycle on Ω^h in the FMG scheme. Estimate the cost of these extra V-cycles. Assume that the cost of a V-cycle on grid Ω^{ph} is proportional to the number of points in that grid, where $p = 2, 4, 8, \ldots, n/2$. Assume also that $\nu_0 = 1$.