Chapter 5

Some Theory

In the previous chapter, we considered some practical questions concerning the implementation, cost, and performance of multigrid methods. The arguments and experiments of that chapter offer good reason to believe that multigrid methods can be extremely effective. Now we must confront some questions on the theoretical side. The goal of this chapter is to present multigrid in a more formal setting and offer an explanation of why these methods work so well. In the first part of this chapter, we sketch the ideas that underlie the convergence theory of multigrid. In the second section, we present what might be called the subspace picture of multigrid. While the terrain in this chapter may seem a bit more rugged than in previous chapters, the reward is an understanding of why multigrid methods are so remarkably effective.

Variational Properties

We first return to a question left unanswered in previous chapters. In expressing the coarse-grid problem, we wrote $A^{2h}\mathbf{u}^{2h} = \mathbf{f}^{2h}$ and said that A^{2h} is the Ω^{2h} version of the original operator A^h. Our first goal is to define the coarse-grid operator A^{2h} precisely.

The argument that follows assumes we are working with the model problem, $-u''(x) = f(x)$, and the corresponding discrete operator A^h. We adopt the notation that Ω^{ph} represents not only the grid with grid spacing ph, but also the space of vectors defined on that grid. In keeping with our former notation, we let $\mathbf{v}^h$ be a computed approximation to the exact solution $\mathbf{u}^h$. For the purpose of this argument, assume that the error in this approximation, $\mathbf{e}^h = \mathbf{u}^h - \mathbf{v}^h$, lies entirely in the range of interpolation, which will be denoted $\mathcal{R}(I_{2h}^h)$. This means that for some vector $\mathbf{u}^{2h} \in \Omega^{2h}$, $\mathbf{e}^h = I_{2h}^h \mathbf{u}^{2h}$. Therefore, the residual equation on Ω^h may be written

$$A^h \mathbf{e}^h = A^h I_{2h}^h \mathbf{u}^{2h} = \mathbf{r}^h. \tag{5.1}$$

In this equation, A^h acts on a vector that lies entirely in the range of interpolation. How does A^h act on $\mathcal{R}(I_{2h}^h)$? Figure 5.1 gives the answer. An arbitrary vector $\mathbf{u}^{2h} \in \Omega^{2h}$ is shown in Fig. 5.1(a); $I_{2h}^h \mathbf{u}^{2h}$, which is the interpolation of $\mathbf{u}^{2h}$ up to Ω^h, is shown in Fig. 5.1(b); and the effect of A^h acting pointwise on $I_{2h}^h \mathbf{u}^{2h}$ is shown in Fig. 5.1(c). We see that $A^h I_{2h}^h \mathbf{u}^{2h}$ is zero at the odd grid points of

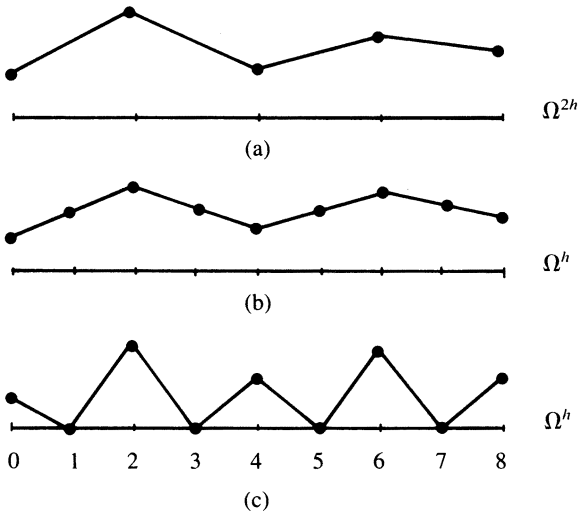

Figure 5.1: *The action of A^h on the range of interpolation $\mathcal{R}(I_{2h}^h)$:* (a) *an arbitrary vector $\mathbf{u}^{2h} \in \Omega^{2h}$;* (b) *the linear interpolant $I_{2h}^h \mathbf{u}^{2h}$; and* (c) *$A^h I_{2h}^h \mathbf{u}^{2h}$, which is zero at the odd grid points on Ω^h.*

Ω^h. The effect is analogous to taking the second derivative of a piecewise linear function.

We may conclude that the odd rows of $A^h I_{2h}^h$ in (5.1) are zero. On the other hand, the even rows of (5.1) correspond to the coarse-grid points of Ω^{2h}. Therefore, we can find a coarse-grid form of the residual equation by dropping the odd rows of (5.1). We do this formally by applying the restriction operator I_h^{2h} to both sides of (5.1). When this is done, the residual equation becomes

$$\underbrace{I_h^{2h} A^h I_{2h}^h}_{A^{2h}} \mathbf{u}^{2h} = I_h^{2h} \mathbf{r}^h.$$

This observation gives us a plausible definition for the coarse-grid operator: $A^{2h} = I_h^{2h} A^h I_{2h}^h$. The terms of A^{2h} may be computed explicitly as shown in Table 5.1. We simply apply $I_h^{2h} A^h I_{2h}^h$ term by term to the jth unit vector $\hat{\mathbf{e}}_j^{2h}$ on Ω^{2h}. This establishes that the jth column of A^{2h} and also, by symmetry, the jth row of A^{2h} are given by

$$\frac{1}{(2h)^2}\begin{pmatrix} -1 & 2 & -1 \end{pmatrix}.$$

We would get the same result if the original problem were simply discretized on Ω^{2h} using the usual second-order finite differences. Therefore, by this definition, A^{2h} really is the Ω^{2h} version of A^h.

The preceding argument was based on the assumption that the error $\mathbf{e}^h$ lies entirely in the range of interpolation. This is not the case in general. If it were, then solving the Ω^{2h} residual equation exactly and doing the two-grid correction would give the exact solution. Nevertheless, the argument does give a sensible definition for A^{2h}. It also leads us to two important properties called the *variational*

	$j-1$		j		$j+1$
$\hat{\mathbf{e}}_j^{2h}$	0		1		0
$I_{2h}^h \hat{\mathbf{e}}_j^{2h}$	0	$\frac{1}{2}$	1	$\frac{1}{2}$	0
$A^h I_{2h}^h \hat{\mathbf{e}}_j^{2h}$	$-\frac{1}{2h^2}$	0	$\frac{1}{h^2}$	0	$-\frac{1}{2h^2}$
$I_h^{2h} A^h I_{2h}^h \hat{\mathbf{e}}_j^{2h}$	$-\frac{1}{4h^2}$		$\frac{1}{2h^2}$		$-\frac{1}{4h^2}$

Table 5.1: *Calculation of the ith row of* $A^{2h} = I_h^{2h} A^h I_{2h}^h$.

properties; they are given by

$$\begin{aligned} A^{2h} &= I_h^{2h} A^h I_{2h}^h \qquad \text{(Galerkin condition)}, \\ I_h^{2h} &= c(I_{2h}^h)^T, \qquad c \in \mathbf{R}. \end{aligned}$$

The first of these, the *Galerkin condition*, is simply the definition of the coarse-grid operator. The second property is the relationship satisfied by the interpolation operator and the full weighting operator defined in Chapter 3. While these properties are not desirable for all applications, they are exhibited by many commonly used operators. They also facilitate the analysis of the two-grid correction scheme. We see a deeper justification of these properties in Chapter 10.

Toward Convergence Theory

Convergence analysis of multigrid methods is difficult and has occupied researchers for several decades. We cannot even pretend to address this problem with the rigor and depth it deserves. Instead, we attempt to give heuristic arguments suggesting that the standard multigrid schemes, when applied to well-behaved problems (for example, scalar elliptic problems), not only work, but work very effectively. Convergence results for these problems have been rigorously proved. For more general problems, new results appear at a fairly steady pace. Where analytical results are lacking, a wealth of computational evidence testifies to the general effectiveness of multigrid methods. Between analysis and experimentation, the multigrid territory is slowly being mapped. However, multigrid convergence analysis is still an open area of computational mathematics.

We begin with a heuristic argument that captures the spirit of rigorous convergence proofs. As we have seen, the smoothing rate (the convergence factor for the oscillatory modes) for the standard relaxation schemes is small and independent of the grid spacing h. Recall that the smooth error modes, which remain after relaxation on one grid, appear more oscillatory on the coarser grids. Therefore, by moving to successively coarser grids, all of the error components on the original fine grid eventually appear oscillatory and are reduced by relaxation. It then follows

that the overall convergence factor for a good multigrid scheme should be small and independent of h.

Now we can be a bit more precise. Denote the original continuous problem (for example, one of our model boundary value problems) by $Au = f$. The associated discrete problem on the fine grid Ω^h is denoted by $A^h\mathbf{u}^h = \mathbf{f}^h$. As before, we let $\mathbf{v}^h$ be an approximation to $\mathbf{u}^h$ on Ω^h. The *discretization error* is defined by

$$E_i^h = u(x_i) - u_i^h, \qquad 1 \le i \le n-1.$$

The discretization error measures how well the exact solution of the discrete problem approximates the exact solution of the original continuous problem. It may be bounded in the discrete L^2 norm in the form

$$\|\mathbf{E}^h\|_h \le Kh^p, \tag{5.2}$$

where K is a positive constant and p is a positive integer. For the model problems in Chapter 1, in which second-order finite differences were used, we have $p = 2$ (see Exercise 11 for a careful derivation of this fact).

Unfortunately, we can seldom solve the discrete problem exactly. The quantity that we have been calling the error, $\mathbf{e}^h = \mathbf{u}^h - \mathbf{v}^h$, will now be called the *algebraic error* to avoid confusion with the discretization error. The algebraic error, as we have seen, measures how well our approximations (generated by relaxation or multigrid) agree with the exact discrete solution.

The purpose of a typical calculation is to produce approximations $\mathbf{v}^h$ that agree with the exact solution of the *continuous problem* $\mathbf{u}$. Let us specify a tolerance ϵ and an error condition such as

$$\|\mathbf{u} - \mathbf{v}^h\|_h < \epsilon, \tag{5.3}$$

where $\mathbf{u} = (u(x_1), \ldots, u(x_{n-1}))^T$ is the vector of exact solution values sampled at the grid points. This condition can be satisfied if we guarantee that both the discretization and algebraic errors are small. Suppose, for example, that

$$\|\mathbf{E}^h\|_h + \|\mathbf{e}^h\|_h < \epsilon.$$

Then, using the triangle inequality,

$$\|\mathbf{u} - \mathbf{v}^h\|_h \le \|\mathbf{u} - \mathbf{u}^h\|_h + \|\mathbf{u}^h - \mathbf{v}^h\|_h = \|\mathbf{E}^h\|_h + \|\mathbf{e}^h\|_h < \epsilon.$$

One way to ensure that $\|\mathbf{E}^h\|_h + \|\mathbf{e}^h\|_h < \epsilon$ is to require that $\|\mathbf{E}^h\|_h < \frac{\epsilon}{2}$ and $\|\mathbf{e}^h\|_h < \frac{\epsilon}{2}$ individually. The first condition determines the grid spacing on the finest grid. Using (5.2), it suggests that we choose

$$h < h^* \equiv \left(\frac{\epsilon}{2K}\right)^{1/p}.$$

The second condition determines how well $\mathbf{v}^h$ must approximate the exact discrete solution $\mathbf{u}^h$. If relaxation or multigrid cycles have been performed until the condition $\|\mathbf{e}^h\| < \frac{\epsilon}{2}$ is met on grid Ω^h, where $h < h^*$, then we have converged to the level of discretization error. In summary, the discretization error determines the critical grid spacing h^*; so (5.3) will be satisfied provided we converge to the level of discretization error on a grid with $h < h^*$.

Consider first a V-cycle scheme applied to a d-dimensional problem with (about) n^d unknowns and $h = \frac{1}{n}$. We assume (and can generally show rigorously) that with fixed cycling parameters, ν_1 and ν_2, the V-cycle scheme has a convergence factor bound, γ, that is independent of h. This V-cycle scheme must reduce the algebraic error from $O(1)$ (the error in the zero initial guess) to $O(h^p) = O(n^{-p})$ (the order of the discretization error). Therefore, the number of V-cycles required, ν, must satisfy $\gamma^\nu = O(n^p)$ or $\nu = O(\log n)$. Because the cost of a single V-cycle is $O(n^d)$, the cost of converging to the level of discretization error with a V-cycle scheme is $O(n^d \log n)$. This is comparable to the computational cost of the best fast direct solvers applied to the model problem.

The FMG scheme costs a little more per cycle than the V-cycle scheme. However, a properly designed FMG scheme can be much more effective overall because it supplies a very good initial guess to the final V-cycles on Ω^h. The key observation in the FMG argument is that before the Ω^h problem is even touched, the Ω^{2h} problem has already been solved to the level of discretization error. This is because of nested iteration, which is designed to provide a good initial guess for the next finer grid. Our goal is to determine how much the algebraic error needs to be reduced by V-cycles on Ω^h. The argument is brief and worthwhile; however, it requires a new tool.

Energy Norms, Inner Products, and Orthogonality. Energy norms and inner products are useful tools in the analysis of multigrid methods. They involve a slight modification of the inner product and norms that we have already encountered. Suppose A is an $n \times n$ symmetric positive definite matrix. Define the *A-inner product* of two vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ by

$$(\mathbf{u}, \mathbf{v})_A \equiv (A\mathbf{u}, \mathbf{v}),$$

where $(\cdot, \cdot)$ is the usual Euclidean inner product on $\mathbf{R}^n$. The *A-norm* now follows in a natural way. Just as $\|\mathbf{u}\| = (\mathbf{u}, \mathbf{u})^{\frac{1}{2}}$, the A-norm is given by

$$\|\mathbf{u}\|_A = (\mathbf{u}, \mathbf{u})_A^{\frac{1}{2}}.$$

The A-norm and A-inner product are sometimes called the *energy* norm and inner product. We can also use the A-inner product to define a new orthogonality relationship. Extending the usual notion of vectors and subspaces, two vectors u and v are *A-orthogonal* if $(u, v)_A = 0$, and we write $u \perp_A v$. We then say that two subspaces U and V are A-orthogonal if, for all $u \in U$ and $v \in V$, we have $u \perp_A v$. In this case, we write $U \perp_A V$.

Our FMG argument can be made in any norm, but it is simplest in the A^h-norm. The goal is to show that one properly designed FMG cycle is enough to ensure that the final algebraic error on Ω^h is below the level of discretization error; that is,

$$\|\mathbf{e}^h\|_{A^h} \leq Kh^p, \tag{5.4}$$

where K is a positive constant that depends on the smoothness of the solution and p is a positive integer. The values of K and p also depend on the norm used to measure the error, so they will generally be different from the constants in (5.2).

The argument is inductive in nature. If Ω^h is the coarsest grid, then FMG is exact and (5.4) is clearly satisfied. Assume now that the Ω^{2h} problem has been solved to the level of discretization error, so that

$$\|\mathbf{e}^{2h}\|_{A^{2h}} \leq K(2h)^p. \tag{5.5}$$

We now use (5.5) to prove (5.4).

The initial algebraic error on Ω^h, before the V-cycles begin, is the difference between the exact fine-grid solution, $\mathbf{u}^h$, and the coarse-grid approximation interpolated to the fine grid:

$$\mathbf{e}_0^h = \mathbf{u}^h - I_{2h}^h \mathbf{v}^{2h}.$$

To estimate the size of this initial error, we must account for the error that might be introduced by interpolation from Ω^{2h} to Ω^h. To do this, we assume that the error in interpolation has the same order (same value of p) as the discretization error and satisfies

$$\|\mathbf{u}^h - I_{2h}^h \mathbf{u}^{2h}\|_{A^h} \leq K\alpha h^p, \tag{5.6}$$

where α is a positive constant. (This sort of bound can be determined rigorously; in fact, α is typically $1 + 2^p$.) A bound for the initial error now follows from the triangle inequality:

$$\begin{aligned} \|\mathbf{e}_0^h\|_{A^h} &= \|\mathbf{u}^h - I_{2h}^h \mathbf{v}^{2h}\|_{A^h} \\ &\leq \|\mathbf{u}^h - I_{2h}^h \mathbf{u}^{2h}\|_{A^h} + \|I_{2h}^h \mathbf{u}^{2h} - I_{2h}^h \mathbf{v}^{2h}\|_{A^h} \quad \text{(triangle inequality)} \\ &= \underbrace{\|\mathbf{u}^h - I_{2h}^h \mathbf{u}^{2h}\|_{A^h}}_{\leq K\alpha h^p} + \underbrace{\|\mathbf{u}^{2h} - \mathbf{v}^{2h}\|_{A^{2h}}}_{\leq K(2h)^p} \quad \text{(Galerkin conditions).} \end{aligned}$$

As indicated, we use (5.5), (5.6), and Exercise 2 to form the following estimate for the norm of the initial error:

$$\|\mathbf{e}_0^h\|_{A^h} \leq K\alpha h^p + K(2h)^p = K(\alpha + 2^p)h^p.$$

To satisfy (5.4), the algebraic error must be reduced from roughly $K(\alpha + 2^p)h^p$ to Kh^p. This means we should use enough V-cycles on Ω^h to reduce the algebraic error by a factor of

$$\beta = \alpha + 2^p.$$

This reduction requires ν V-cycles, where $\gamma^\nu \leq \beta$. Because β is $O(1)$, it follows that ν is also $O(1)$. (Typically, $\beta = 5$ or 9 and $\gamma \approx 0.1$ for a V(2,1)-cycle, so $\nu = 1$.) In other words, the preliminary cycling through coarser grids gives such a good initial guess that only $O(1)$ V-cycles are needed on the finest grid. This means that the total computational cost of FMG is $O(n^d)$, which is optimal.

This discussion is meant to give some feeling for the rigorous arguments that can be used to establish the convergence properties of the basic multigrid algorithms. These results cannot be pursued much further at this point and must be left to the multigrid literature. Instead, we turn to a different perspective on why multigrid works.

Spectral and Algebraic Pictures

Much of this section is devoted to an analysis of the two-grid correction scheme, which lies at the heart of multigrid. Recall that the V-cycle is just nested applications of the two-grid correction scheme and that the FMG method is just

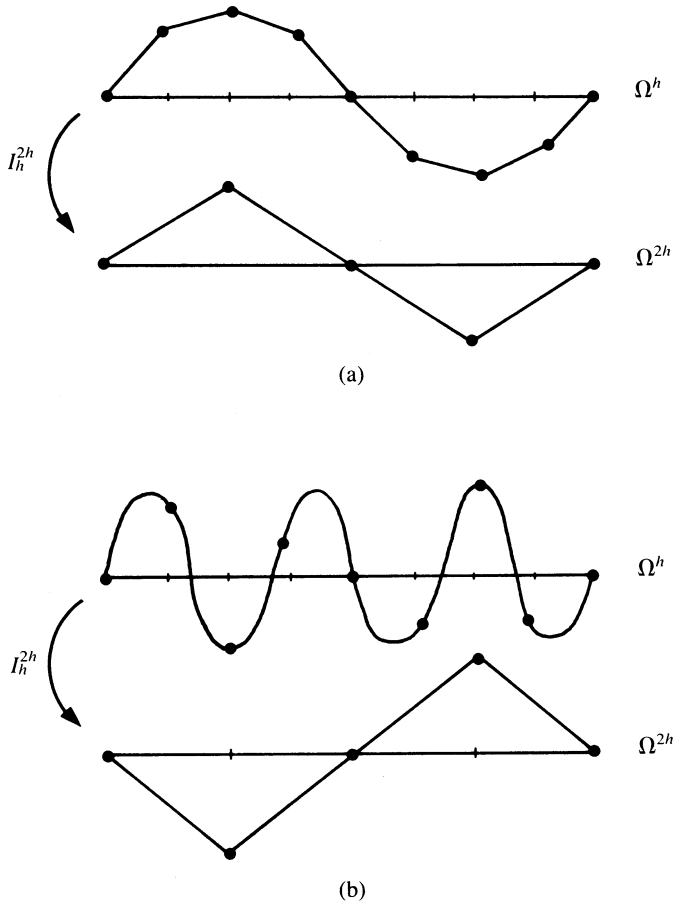

Figure 5.2: *The full weighting operator I_h^{2h} acting on* (a) *a smooth mode of Ω^h ($k = 2$ and $n = 8$) and* (b) *an oscillatory mode of Ω^h ($k = 6$ and $n = 8$). In the first case, the result is a multiple of the $k = 2$ mode on Ω^{2h}. In the second case, the result is a multiple of the $n - k = 2$ mode on Ω^{2h}.*

repeated applications of the V-cycle on various grids. Therefore, an understanding of the two-grid correction scheme is essential for a complete explanation of the basic multigrid methods.

We begin with a detailed look at the intergrid transfer operators. Consider first the full weighting (restriction) operator, I_h^{2h}. Recall that I_h^{2h} maps $\mathbf{R}^{n-1} \rightarrow \mathbf{R}^{\frac{n}{2}-1}$. It has rank $\frac{n}{2} - 1$ and null space $N(I_h^{2h})$ of dimension $\frac{n}{2}$. It is important to understand what we call the spectral properties of I_h^{2h}. In particular, how does I_h^{2h} act upon the modes of the original operator A^h?

Recall that the modes of A^h for the one-dimensional model problem are given by

$$w_{k,j}^h = \sin\left(\frac{jk\pi}{n}\right), \qquad 1 \leq k \leq n-1, \quad 0 \leq j \leq n.$$

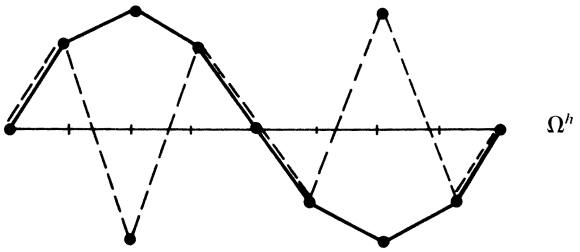

Figure 5.3: *A pair of complementary modes on a grid with $n = 8$ points. The solid line shows the $k = 2$ mode. The dashed line shows the $k' = n - k = 6$ mode.*

The full weighting operator may be applied directly to these vectors. The result of I_h^{2h} acting on the smooth modes is (Exercise 4)

$$I_h^{2h}\mathbf{w}_k^h = \cos^2\left(\frac{k\pi}{2n}\right)\mathbf{w}_k^{2h}, \qquad 1 \le k \le \frac{n}{2}\,.$$

This says that I_h^{2h} acting on the kth (smooth) mode of A^h produces a constant times the kth mode of A^{2h} when $1 \le k \le \frac{n}{2}$. This property is illustrated in Fig. 5.2(a). For the oscillatory modes, with $\frac{n}{2} < k < n-1$, we have (Exercise 5)

$$I_h^{2h}\mathbf{w}_{k'}^h = -\sin^2\left(\frac{k\pi}{2n}\right)\mathbf{w}_k^{2h}, \qquad 1 \le k < \frac{n}{2}\,,$$

where $k' = n-k$. This says that I_h^{2h} acting on the $(n-k)$th mode of A^h produces a constant multiple of the kth mode of A^{2h}. This property, illustrated in Fig. 5.2(b), is similar to the aliasing phenomenon discussed earlier. The oscillatory modes on Ω^h cannot be represented on Ω^{2h}. As a result, the full weighting operator transforms these modes into relatively smooth modes on Ω^{2h}.

In summary, we see that both the kth and $(n-k)$th modes on Ω^h become the kth mode on Ω^{2h} under the action of full weighting. We refer to this pair of fine-grid modes $\{\mathbf{w}_k^h, \mathbf{w}_{n-k}^h\}$ as *complementary modes*. Letting $W_k^h = \text{span}\{\mathbf{w}_k^h, \mathbf{w}_{n-k}^h\}$, we have that

$$I_h^{2h} : W_k^h \to \text{span}\{\mathbf{w}_k^{2h}\}.$$

In passing, it is interesting to note the relationship between two complementary modes. It may be shown (Exercise 6) that $w_{n-k,j}^h = (-1)^{j+1} w_{k,j}^h$. Figure 5.3 illustrates the smooth and oscillatory nature of a pair of complementary modes.

As noted earlier, the full weighting operator has a nontrivial null space, $N(I_h^{2h})$. It may be verified (Exercise 7) that this subspace is spanned by the vectors $\mathbf{n}_j = A^h\hat{\mathbf{e}}_j^h$, where j is odd and $\hat{\mathbf{e}}_j^h$ is the jth unit vector on Ω^h. As shown in Fig. 5.4, the basis vectors $\mathbf{n}_j$ appear oscillatory. However, they do not coincide with the oscillatory modes of A^h. In fact, an expansion of $\mathbf{n}_j$ in terms of the modes of A^h requires all of the modes. Thus, the null space of I_h^{2h} possesses both smooth and oscillatory modes of A^h.

Having established the necessary properties of the full weighting operator, we now examine the interpolation operator I_{2h}^h in the same way. Recall that I_{2h}^h maps

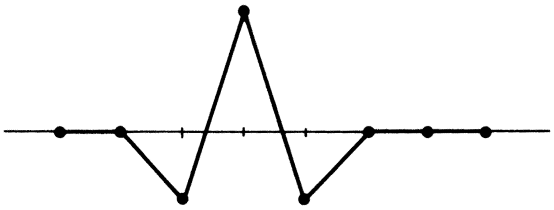

Figure 5.4: *A typical basis vector of the null space of the full weighting operator* $N(I_h^{2h})$.

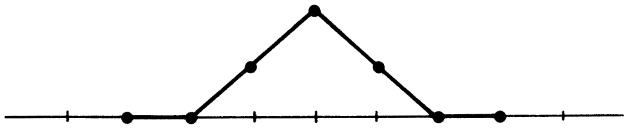

Figure 5.5: *A typical basis vector of the range of interpolation* $\mathcal{R}(I_{2h}^h)$.

$\mathbf{R}^{\frac{n}{2}-1} \rightarrow \mathbf{R}^{n-1}$ and has full rank. In order to establish the spectral properties of I_{2h}^h, we ask how I_{2h}^h acts on the modes of A^{2h}. Letting

$$w_{k,j}^{2h} = \sin\left(\frac{jk\pi}{n/2}\right), \qquad 1 \leq k < \frac{n}{2}, \quad 0 \leq j \leq \frac{n}{2},$$

be the Ω^{2h} modes, we can show (Exercise 8) that I_{2h}^h does not preserve these modes. The calculation shows that

$$I_{2h}^h \mathbf{w}_k^{2h} = c_k \mathbf{w}_k^h - s_k \mathbf{w}_{k'}^h, \qquad 1 \leq k < \frac{n}{2}, \quad k' = n - k,$$

where $c_k = \cos^2\left(\frac{k\pi}{2n}\right)$ and $s_k = \sin^2\left(\frac{k\pi}{2n}\right)$. We see that I_{2h}^h acting on the kth mode of Ω^{2h} produces not only the kth mode of Ω^h but also the complementary mode $\mathbf{w}_{k'}^h$. This fact exposes the interesting property that interpolation of smooth modes on Ω^{2h} excites (to some degree) oscillatory modes on Ω^h. It should be noted that for a very smooth mode on Ω^h with $k \ll n/2$,

$$I_{2h}^h \mathbf{w}_k^{2h} = \left[1 - O\left(\frac{k^2}{n^2}\right)\right] \mathbf{w}_k^h + O\left(\frac{k^2}{n^2}\right) \mathbf{w}_{k'}^h.$$

In this case, the result of interpolation is largely the corresponding smooth mode on Ω^h with very little contamination from the complementary oscillatory mode. As it has been defined, I_{2h}^h is a *second-order* interpolation operator because the magnitude of the spurious oscillatory mode is $O(\frac{k^2}{n^2})$.

We have already anticipated the importance of the range of interpolation, $\mathcal{R}(I_{2h}^h)$. A basis for $\mathcal{R}(I_{2h}^h)$ is given by the columns of I_{2h}^h. While these basis vectors appear smooth, as Fig. 5.5 shows, they do not coincide with the smooth modes of A^h. In fact, it may be shown that any one of these basis vectors requires all modes of A^h for a full representation. In other words, the range of interpolation contains both smooth and oscillatory modes of A^h.

With this investigation of the intergrid transfer operators, we now return to the two-grid correction scheme. We begin with an observation made in Chapter 2 that a stationary linear iteration may be expressed in the form

$$\mathbf{v}^{(1)} = (I - BA)\mathbf{v}^{(0)} + B\mathbf{f} = R\mathbf{v}^{(0)} + B\mathbf{f},$$

where B is a specified matrix and $R = I - BA$ is the iteration matrix for the method. It follows that m sweeps of the iteration can be represented by

$$\mathbf{v}^{(m)} = R^m\mathbf{v}^{(0)} + C(\mathbf{f}),$$

where $C(\mathbf{f})$ represents a series of operations on $\mathbf{f}$.

We can now turn to the two-grid correction scheme. The steps of this scheme, with an exact solution on the coarse grid, are given by the following procedure:

- Relax ν times on Ω^h with scheme R: $\mathbf{v}^h \leftarrow R^\nu\mathbf{v}^h + C(\mathbf{f})$.
- Full weight $\mathbf{r}^h$ to Ω^{2h}: $\mathbf{f}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h)$.
- Solve the residual equation exactly: $\mathbf{v}^{2h} = (A^{2h})^{-1}\mathbf{f}^{2h}$.
- Correct the approximation on Ω^h: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{v}^{2h}$.

If we now take this process one step at a time, it may be represented in terms of a single replacement operation:

$$\mathbf{v}^h \leftarrow R^\nu\mathbf{v}^h + C(\mathbf{f}) + I_{2h}^h(A^{2h})^{-1}I_h^{2h}(\mathbf{f}^h - A^h(R^\nu\mathbf{v}^h + C(\mathbf{f}))).$$

The exact solution $\mathbf{u}^h$ is unchanged by the two-grid correction scheme. Therefore,

$$\mathbf{u}^h = R^\nu\mathbf{u}^h + C(\mathbf{f}) + I_{2h}^h(A^{2h})^{-1}I_h^{2h}(\mathbf{f}^h - A^h(R^\nu\mathbf{u}^h + C(\mathbf{f}))).$$

By subtracting these last two expressions, we can see how the two-grid correction operator, which we now denote TG, acts upon the error, $\mathbf{e}^h = \mathbf{u}^h - \mathbf{v}^h$. We find that

$$\mathbf{e}^h \leftarrow \left[I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\right]R^\nu\mathbf{e}^h \equiv TG\mathbf{e}^h. \tag{5.7}$$

As in Chapter 2, we imagine that the error can be expressed as a linear combination of the modes of A^h. This leads us to ask how TG acts upon the modes of A^h. However, TG consists of R, A^h, $(A^{2h})^{-1}$, I_h^{2h}, and I_{2h}^h, and we now know how each of these operators acts upon the modes of A^h. For the moment, consider the two-grid correction scheme TG with no relaxation ($\nu = 0$). Using all of the spectral properties we have just discovered, it may be shown (Exercise 9) that the coarse-grid correction operator, TG, is invariant on the subspaces $W_k^h = \text{span}\{\mathbf{w}_k^h, \mathbf{w}_{k'}^h\}$; that is,

$$TG\mathbf{w}_k = s_k\mathbf{w}_k + s_k\mathbf{w}_{k'}, \tag{5.8}$$

$$TG\mathbf{w}_{k'} = c_k\mathbf{w}_k + c_k\mathbf{w}_{k'}, \quad 1 \le k \le \frac{n}{2}, \quad k' = n - k, \tag{5.9}$$

where $c_k = \cos^2\left(\frac{\pi k}{2n}\right)$ and $s_k = \sin^2\left(\frac{\pi k}{2n}\right)$.

This implies that when TG is applied to a smooth or oscillatory mode, the same mode and its complement result. But it is important to look at the amplitudes of

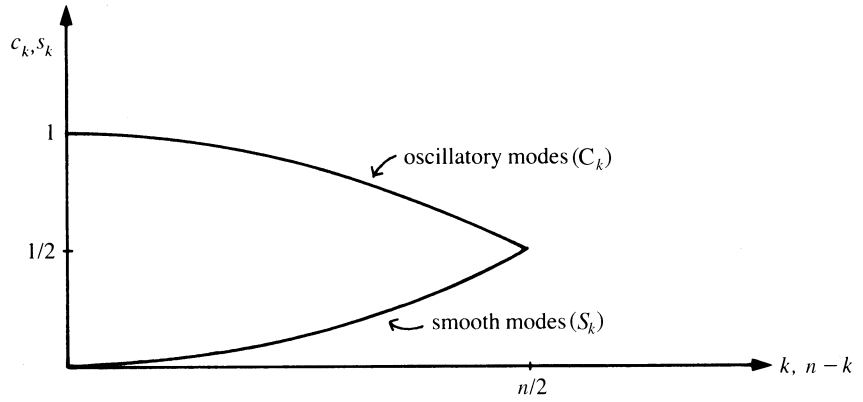

Figure 5.6: *Damping factor for the two-grid correction operator TG, without relaxation, acting on the oscillatory modes with wavenumbers $n-k$ (upper curve) and on the smooth modes with wavenumbers k (lower curve) for $1 \le k \le \frac{n}{2}$.*

the resulting modes. Consider the case of TG acting on very smooth modes and very oscillatory modes with $k \ll n$. Then (5.8) and (5.9) become

$$\begin{aligned} TG\mathbf{w}_k &= O\left(\frac{k^2}{n^2}\right)\mathbf{w}_k + O\left(\frac{k^2}{n^2}\right)\mathbf{w}_{k'}, \\ TG\mathbf{w}_{k'} &= \left[1 - O\left(\frac{k^2}{n^2}\right)\right]\mathbf{w}_k + \left[1 - O\left(\frac{k^2}{n^2}\right)\right]\mathbf{w}_{k'}, \quad 1 \le k \le \frac{n}{2}, \quad k' = n-k. \end{aligned}$$

TG acting on smooth modes produces smooth and oscillatory modes with very small amplitudes. Therefore, the two-grid correction scheme is effective at eliminating smooth components of the error. However, when TG acts upon highly oscillatory modes, it produces smooth and oscillatory modes with $O(1)$ amplitudes. Therefore, two-grid correction, without relaxation, does not damp oscillatory modes. Figure 5.6 illustrates this behavior of the two-grid correction scheme with no relaxation by showing the damping factors, c_k and s_k, for the smooth and oscillatory components of the error.

We now bring relaxation into the picture. Knowing its spectral properties, we can anticipate that relaxation will balance perfectly the action of TG without relaxation. We now include ν steps of a relaxation method R and assume for simplicity that R does not mix the modes of A^h. Many other relaxation methods can be analyzed without this assumption. As before, let λ_k be the eigenvalue of R associated with the kth mode $\mathbf{w}_k$. Combining all of these observations with (5.7), the action of TG *with* relaxation is given by (Exercise 10)

$$TG\mathbf{w}_k = \lambda_k^\nu s_k \mathbf{w}_k + \lambda_k^\nu s_k \mathbf{w}_{k'}, \tag{5.10}$$

$$TG\mathbf{w}_{k'} = \lambda_{k'}^\nu c_k \mathbf{w}_k + \lambda_{k'}^\nu c_k \mathbf{w}_{k'}, \quad 1 \le k \le \frac{n}{2},\ k' = n-k. \tag{5.11}$$

We know that the smoothing property of relaxation has the strongest effect on the oscillatory modes. This is reflected in the term $\lambda_{k'}^\nu$, which is small. At the same time, the two-grid correction scheme alone (without relaxation) eliminates

the smooth modes. This is reflected in the s_k terms. Thus, all terms of (5.10) and (5.11) are small, particularly for $k \ll \frac{n}{2}$ or as ν becomes large. The result is a complete process in that both smooth and oscillatory modes of the error are well damped.

We have now completed what we call the spectral picture of multigrid. By examining how various operators act on the modes of A^h, we have determined the effect of the entire two-grid correction operator on those modes. This analysis explains how the two-grid correction process eliminates both the smooth and oscillatory components of the error.

Solvability and the Fundamental Theorem of Linear Algebra. Suppose we have a matrix $A \in \mathbf{R}^{m \times n}$. The fundamental theorem of linear algebra states that the range (column space) of the matrix, $\mathcal{R}(A)$, is equal to the orthogonal complement of $\mathcal{N}(\mathcal{A}^{\mathcal{T}})$, the null space of A^T. Thus, spaces $\mathbf{R}^m$ and $\mathbf{R}^n$ can be orthogonally decomposed as follows:

$$\begin{aligned} \mathbf{R}^m &= \mathcal{R}(A) \oplus \mathcal{N}(A^T), \\ \mathbf{R}^n &= \mathcal{R}(A^T) \oplus \mathcal{N}(A). \end{aligned}$$

For the equation $A\mathbf{x} = \mathbf{f}$ to have a solution, it is necessary that the vector $\mathbf{f}$ lie in $\mathcal{R}(A)$. Thus, an equivalent condition is that $\mathbf{f}$ be orthogonal to every vector in $\mathcal{N}(A^T)$. For the equation $A\mathbf{x} = \mathbf{f}$ to have a *unique* solution, it is necessary that $\mathcal{N}(A) = \{\mathbf{0}\}$. Otherwise, if $\mathbf{x}$ is a solution and $\mathbf{y} \in \mathcal{N}(A)$, then $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{f} + \mathbf{0} = \mathbf{f}$, so the solution $\mathbf{x}$ is not unique.

There is another vantage point from which to view the coarse-grid correction scheme. This perspective will lead to what we will call the algebraic picture of multigrid. With both the spectral and the algebraic picture before us, it will be possible to give a good qualitative explanation of multigrid. Let us now look at the algebraic structure of the two-grid correction scheme.

The variational properties introduced earlier now become important. Recall that these properties are given by

$$\begin{aligned} A^{2h} &= I_h^{2h} A^h I_{2h}^h \qquad \text{(Galerkin property)}, \\ I_h^{2h} &= c(I_{2h}^h)^T, \qquad c \in \mathbf{R}. \end{aligned}$$

The two-grid correction scheme involves transformations between the space of fine-grid vectors, Ω^h, and the space of coarse-grid vectors, Ω^{2h}. Figure 5.7 diagrams these two spaces and the action of the full weighting and interpolation operators.

As we have already seen, the range of interpolation, $\mathcal{R}(I_{2h}^h)$, and the null space of full weighting, $N(I_h^{2h})$, both reside in Ω^h and have dimensions of roughly $\frac{n}{2}$. From the orthogonality relationships between the subspaces of a linear operator (Fundamental Theorem of Linear Algebra), we know that

$$N(I_h^{2h}) \perp \mathcal{R}[(I_h^{2h})^T].$$

By the second variational property, it then follows that

$$N(I_h^{2h}) \perp \mathcal{R}(I_{2h}^h).$$

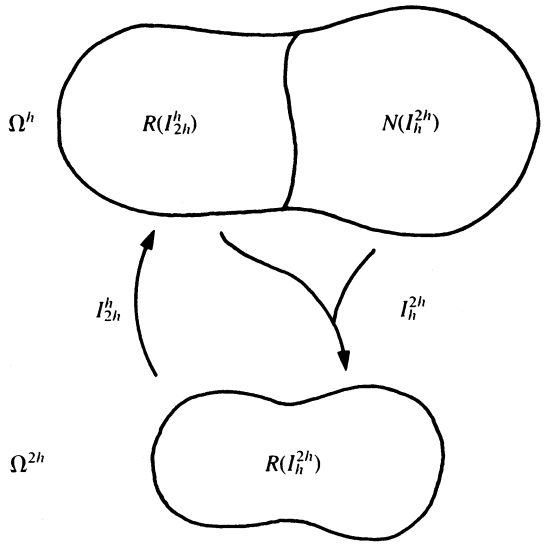

Figure 5.7: *Schematic drawing of the space of fine-grid vectors, Ω^h; the space of coarse-grid vectors, Ω^{2h}; and their subspaces and the intergrid transfer operators, I_{2h}^h and I_h^{2h}.*

The fact that the range of interpolation is orthogonal to the null space of full weighting is significant.

We will now use the notion of A-orthogonality to rewrite the above orthogonality relationship. The fact that $N(I_h^{2h}) \perp \mathcal{R}(I_{2h}^h)$ means that $(\mathbf{q}^h, \mathbf{u}^h) = 0$ whenever $\mathbf{q}^h \in \mathcal{R}(I_{2h}^h)$ and $I_h^{2h}\mathbf{u}^h = \mathbf{0}$. This is equivalent to the condition that $(\mathbf{q}^h, A^h\mathbf{u}^h) = 0$ whenever $\mathbf{q}^h \in \mathcal{R}(I_{2h}^h)$ and $I_h^{2h}A^h\mathbf{u}^h = \mathbf{0}$. This last condition may be written as

$$N(I_h^{2h}A^h) \perp_{A^h} \mathcal{R}(I_{2h}^h);$$

that is, the null space of $I_h^{2h}A^h$ is A^h-orthogonal to the range of interpolation.

This orthogonality property allows the space Ω^h to be decomposed in the form

$$\Omega^h = \mathcal{R}(I_{2h}^h) \oplus N(I_h^{2h}A^h).$$

This means that if $\mathbf{e}^h$ is a vector in Ω^h, then it may always be expressed as

$$\mathbf{e}^h = \mathbf{s}^h + \mathbf{t}^h,$$

where $\mathbf{s}^h \in \mathcal{R}(I_{2h}^h)$ and $\mathbf{t}^h \in N(I_h^{2h}A^h)$.

It will be helpful to interpret the vectors $\mathbf{s}^h$ and $\mathbf{t}^h$. Since $\mathbf{s}^h$ is an element of $\mathcal{R}(I_{2h}^h)$, it must satisfy $\mathbf{s}^h = I_{2h}^h\mathbf{q}^{2h}$, where $\mathbf{q}^{2h}$ is some vector of Ω^{2h}. We observed the smoothing effect of interpolation and noted the smooth appearance of the basis vectors of $\mathcal{R}(I_{2h}^h)$. For this reason, we associate $\mathbf{s}^h$ with the smooth components of the error. We also noted the oscillatory appearance of the basis vectors of $N(I_h^{2h})$. For this reason, we associate $\mathbf{t}^h$ with the oscillatory components of the error.

We may now consider the two-grid correction operator in light of these subspace properties. The two-grid correction operator without relaxation is

$$TG = I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h.$$

First note that if $\mathbf{s}^h \in \mathcal{R}(I_{2h}^h)$, then $\mathbf{s}^h = I_{2h}^h \mathbf{q}^{2h}$ for some vector $\mathbf{q}^{2h}$ in Ω^{2h}. We then have that

$$TG\mathbf{s}^h = \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] I_{2h}^h \mathbf{q}^{2h}. \tag{5.12}$$

However, by the first of the variational properties, $I_h^{2h} A^h I_{2h}^h = A^{2h}$. Therefore, $TG\mathbf{s}^h = \mathbf{0}$. This gives us the important result that any vector in the range of interpolation also lies in the null space of the two-grid correction operator, that is,

$$N(TG) \supset \mathcal{R}(I_{2h}^h).$$

Having seen how TG acts on $\mathcal{R}(I_{2h}^h)$, now consider a vector $\mathbf{t}^h$ in $N(I_h^{2h} A^h)$. This case is even simpler. We have

$$TG\mathbf{t}^h = \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] \mathbf{t}^h. \tag{5.13}$$

Because $I_h^{2h} A^h \mathbf{t}^h = \mathbf{0}$, we conclude that $TG\mathbf{t}^h = \mathbf{t}^h$. This says that TG is the identity when it acts on $N(I_h^{2h} A^h)$. This implies that the dimension of $N(TG)$ cannot exceed the dimension of $\mathcal{R}(I_{2h}^h)$. We therefore have the stronger result (Exercise 11) that

$$N(TG) = \mathcal{R}(I_{2h}^h).$$

This argument gives us the effect of the two-grid correction operator on the two orthogonal subspaces of Ω^h. Let us now put these algebraic results together with the spectral picture. We have established two independent ways to decompose the space of fine-grid vectors, Ω^h. We have the spectral decomposition

$$\Omega^h = L \oplus H = \left\{ \begin{array}{c} \text{Low-frequency modes} \\ 1 \le k < \dfrac{n}{2} \end{array} \right\} \oplus \left\{ \begin{array}{c} \text{High-frequency modes} \\ \dfrac{n}{2} \le k < n \end{array} \right\}$$

and the subspace decomposition

$$\Omega^h = \mathcal{R}(I_{2h}^h) \oplus N(I_h^{2h} A^h).$$

We can now give a schematic illustration of the two-grid correction scheme as it works on an arbitrary error vector. The diagrams in Figs. 5.8–5.11 are rather unconventional and may require some deliberation. However, they do incorporate all of the spectral and subspace results elaborated on in this chapter.

We first focus on the upper diagrams of each figure. These diagrams portray the space Ω^h as the plane of the page. As we have seen, Ω^h may be decomposed in two ways. These two decompositions are represented by two pairs of orthogonal axes labeled (H, L) for the high-frequency/low-frequency decomposition and (N, R) for the $N(I_h^{2h} A^h)/\mathcal{R}(I_{2h}^h)$ decomposition. The L and R axes are more nearly aligned, suggesting that the smooth, low-frequency modes are associated with the range of interpolation.

Initially, an arbitrary error vector $\mathbf{e}^h$ in Ω^h appears as a point in the plane as shown in Fig. 5.8. This vector has projections on all four axes. Specifically, consider the projections on the R and N axes, which we have called $\mathbf{s}^h$ and $\mathbf{t}^h$, respectively. Both $\mathbf{s}^h$ and $\mathbf{t}^h$ may be further projected onto the L and H axes. The projections of $\mathbf{s}^h$ on the L and H axes are denoted $\mathbf{s}_L$ and $\mathbf{s}_H$, respectively. The projections of $\mathbf{t}^h$ on the L and H axes are denoted $\mathbf{t}_L$ and $\mathbf{t}_H$, respectively. This gives the picture for the initial error, before any relaxation is done.

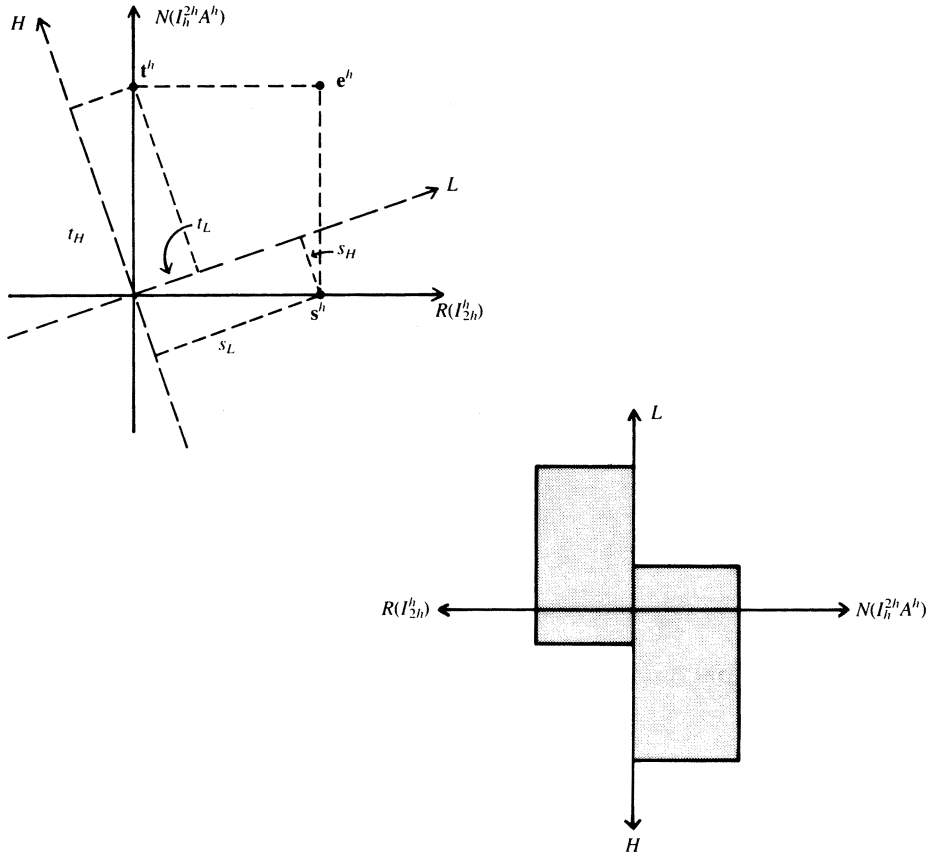

Figure 5.8: *The top diagram shows the space Ω^h decomposed along the (low-frequency, high-frequency) axes and along the $(\mathcal{R}(I_{2h}^h), N(I_h^{2h}A^h))$ axes. The initial error, $\mathbf{e}^h$, appears as a point with projections on all four axes. The lower diagram gives a schematic "energy budget" of the initial error. The error energy is divided between $\mathcal{R}(I_{2h}^h)$ and $N(I_h^{2h}A^h)$ and further divided between low- and high-frequency modes.*

The upper diagram in Fig. 5.9 shows the effect of several relaxation sweeps on the fine grid. Assume that enough relaxation sweeps are done to eliminate entirely the high-frequency components of the error. In the diagram, the resulting error has no component along the H-axis and the point representing $\mathbf{e}^h$ is projected down to the L-axis. The diagram also shows the new projections on the R and N axes. Notice that the component of $\mathbf{e}^h$ in $\mathcal{R}(I_{2h}^h)$ has actually increased.

The upper diagram of Fig. 5.10 shows the effect of the rest of the two-grid correction scheme. Since $\mathcal{R}(I_{2h}^h)$ is the null space of TG, the component of $\mathbf{e}^h$ along the R-axis vanishes. Therefore, this step is represented by a projection directly onto the N-axis. This is consistent with our observation that TG acts as the identity on $N(I_h^{2h}A^h)$. This new projection has eliminated a large amount of the initial error. However, as we proved earlier, the coarse-grid correction operator (without

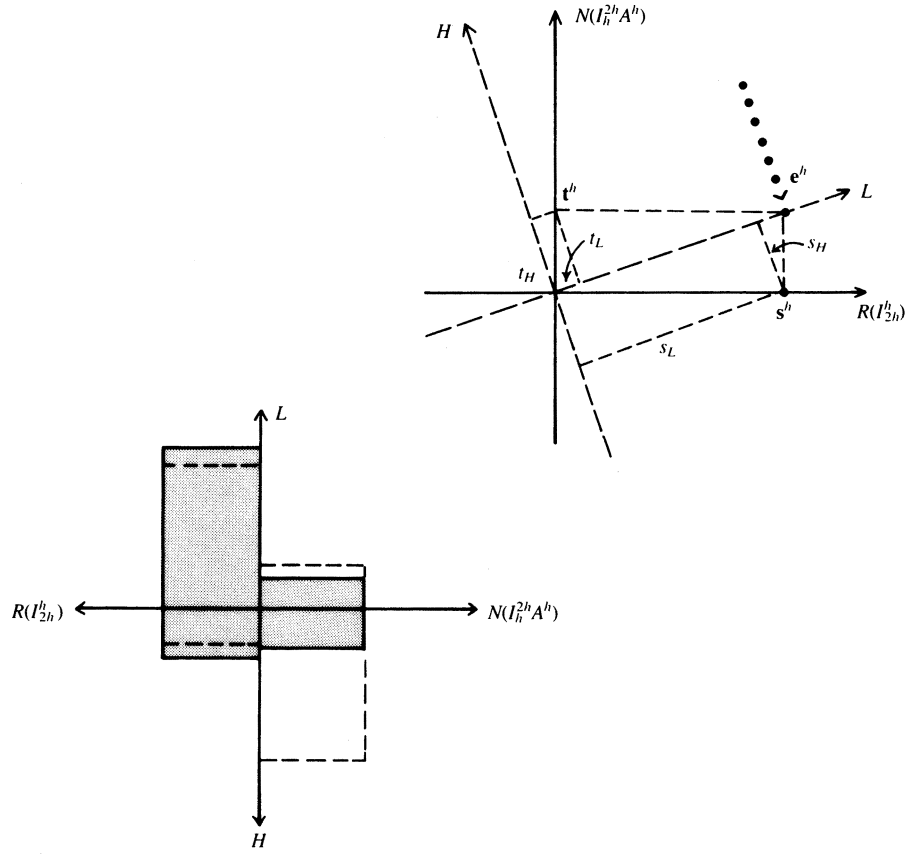

Figure 5.9: *The top diagram shows the effect of several fine-grid relaxation sweeps. Assuming that all oscillatory components are eliminated, the error is projected onto the low-frequency axis (L). The lower diagram shows the corresponding changes in the energy budget. Dashed lines indicate the previous energy budget. Shaded areas show the current budget.*

relaxation) does excite oscillatory modes, which can be seen in the diagram as a nonzero component in the H direction.

The upper diagram of Fig. 5.11 shows the error after more relaxation sweeps on the fine grid. Assuming again that relaxation eliminates all high-frequency error modes, this relaxation sweep is represented as a projection directly onto the L-axis. This further reduces a significant amount of the error.

The pattern should now be evident. By combining relaxation with corrections from the coarse-grid residual equation, we alternately project onto the L and N axes. In doing this, the error vector is driven toward the origin (zero error) in a way that is reminiscent of the convergence of a fixed point iteration.

The lower diagrams of Figs. 5.8–5.11 attempt to illustrate the two-grid correction scheme in terms of the "energy budget" of the error. Initially, the error has a certain energy or magnitude. Part of this energy resides in $\mathcal{R}(I_{2h}^h)$, which is drawn

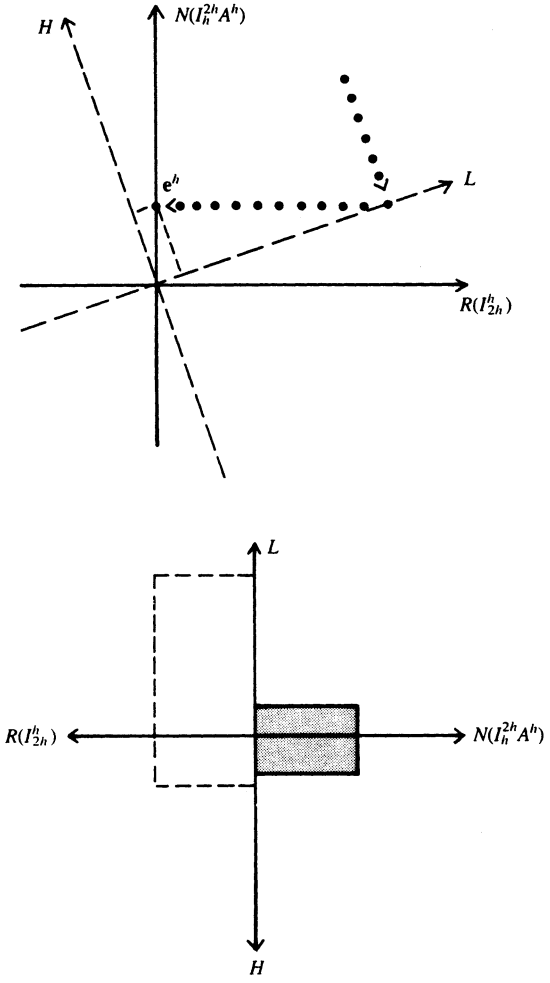

Figure 5.10: *The top diagram shows the error after doing the remainder of the two-grid correction. All components in $\mathcal{R}(I_{2h}^h)$ are eliminated as indicated by the projection onto the N-axis. The lower diagram shows the corresponding changes in the energy budget.*

on the left of the vertical line; the remainder of the energy resides in $N(I_h^{2h}A^h)$, which is drawn on the right of that line. The energy in $\mathcal{R}(I_{2h}^h)$ may be divided further between the low-frequency modes (above the horizontal line) and the high-frequency modes (below the horizontal line). We see that the initial error is fairly evenly divided between $\mathcal{R}(I_{2h}^h)$ and $N(I_h^{2h}A^h)$. As we noted before, most of the error in $\mathcal{R}(I_{2h}^h)$ is associated with the smooth, low-frequency modes; most of the error in $N(I_h^{2h}A^h)$ is associated with the oscillatory, high-frequency modes.

The lower diagram of Fig. 5.9 shows the energy budget after several effective relaxation sweeps. The dotted lines indicate the previous budget status. The shaded regions indicate the budget after relaxation. As described above, relaxation eliminates the high-frequency error components, but at the same time increases the component of the error in $\mathcal{R}(I_{2h}^h)$.

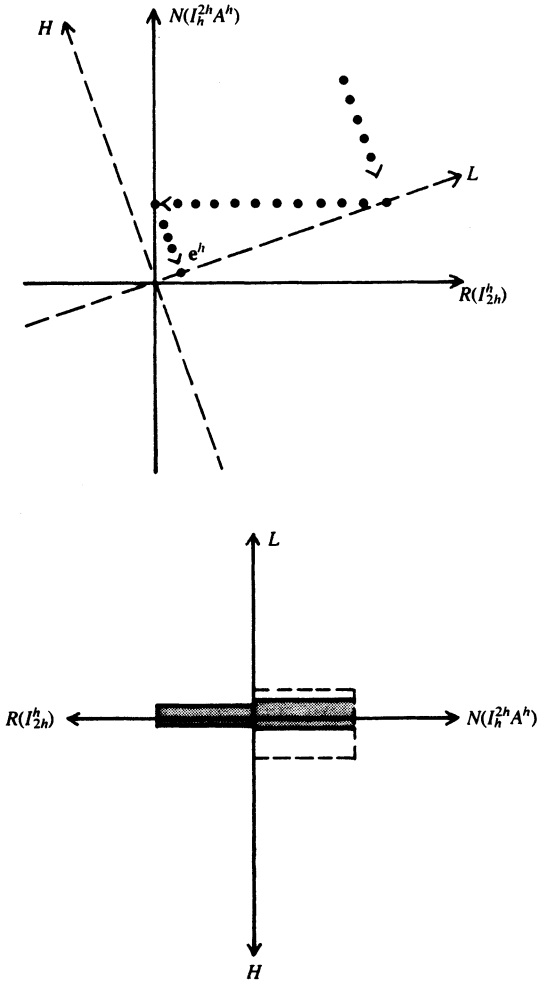

Figure 5.11: *The top diagram shows the effect of several more fine-grid relaxations on the error. Once again, all high-frequency components are removed as indicated by a projection onto the L-axis. The lower diagram shows the corresponding changes in the energy budget.*

The lower diagram of Fig. 5.10 shows the energy budget after a correction step. As we saw earlier, this step eliminates all of the error in $\mathcal{R}(I_{2h}^h)$. Finally, further relaxation (Fig. 5.11) reduces the remaining oscillatory error modes, but reintroduces error with a small magnitude into $\mathcal{R}(I_{2h}^h)$. A continuation of this cycle will further reduce the energy of the error; this reduction signifies convergence.

The subspace iteration diagrams, together with the energy budget diagrams, illustrate the way in which relaxation and correction work together. These two processes complement each other remarkably. When they are applied in tandem, the result is an extremely effective algorithm.

It should be remembered that this entire discussion has dealt only with a two-grid scheme. The V-cycle uses this scheme at all levels, ensuring that relaxation is

always directed at the oscillatory modes of the current grid. As we have seen, two-grid correction without relaxation takes care of the smooth error components on the current fine grid. Therefore, by adding relaxation at all levels, all components of the error are eventually acted upon and quickly removed.

The efficiency of the two-grid scheme is further amplified by the FMG method. This scheme uses nested V-cycles to compute accurate initial guesses on coarser grids before relaxing on the finer grids. This ensures that a particular coarse-grid problem is solved to the level of discretization error before the more expensive fine-grid relaxations are begun. But even in these more elaborate procedures, it is the combination of relaxation and correction that provides the underlying power. Therefore, the arguments of this chapter, although somewhat qualitative, hold the key to the remarkable effectiveness of multigrid.

Exercises

1. **Energy norm.** Assume A is symmetric positive definite. As defined in this chapter, the A-energy inner product and the A-energy norm are given by

$$(\mathbf{u}, \mathbf{v})_A = (A\mathbf{u}, \mathbf{v}) \quad \text{and} \quad \|\mathbf{u}\|_A^2 = (\mathbf{u}, \mathbf{u})_A.$$

   (a) Show that these are acceptable definitions for an inner product and a norm.

   (b) Show that $\|\mathbf{r}\|_2 = \|\mathbf{e}\|_{A^2}$.

   (c) The error norm $\|\mathbf{e}\|_2$ is generally not computable. Is $\|\mathbf{e}\|_A$ computable? Is $\|\mathbf{e}\|_{A^2}$ computable?

2. **FMG error analysis.** A key step in the FMG error analysis is showing that

$$\|I_{2h}^h \mathbf{u}^{2h} - I_{2h}^h \mathbf{v}^{2h}\|_{A^h} = \|\mathbf{u}^{2h} - \mathbf{v}^{2h}\|_{A^{2h}}.$$

   Use the Galerkin property and the property of inner products that

$$(B\mathbf{u}, \mathbf{v}) = (\mathbf{u}, B^T\mathbf{v})$$

   to prove this equality for any two coarse-grid vectors.

3. **Discretization error.** The goal of this exercise is to consider the model problem in one dimension and show carefully that the discrete L^2 norm of the discretization error is bounded by a constant times h^2 (see equation (5.2)). Consider the boundary value problem

$$\begin{aligned} -u''(x) &= f(x) \quad \text{on} \quad (0,1), \\ u(0) = u(1) &= 0. \end{aligned}$$

   The discretized form of the problem is

$$\begin{aligned} (A^h\mathbf{u}^h)_i = -\frac{u_{i+1}^h - 2u_i^h + u_{i-1}^h}{h^2} &= f(x_i), \quad 1 \le i \le n, \\ u_0^h = u_{n+1}^h &= 0. \end{aligned}$$

(a) Let $\mathbf{u}$ be the vector consisting of values of the exact solution u sampled at the grid points; that is, $\mathbf{u}_i = u(x_i)$. Expanding $u(x_{i\pm1})$ in a Taylor series about x_i, show that

$$(A^h\mathbf{u})_i = -u''(x_i) - \frac{h^2}{24}(u^{(iv)}(\xi^+) + u^{(iv)}(\xi^-)),$$

where $x_{i-1} < \xi^- < x_i$ and $x_i < \xi^+ < x_{i+1}$.

(b) Assuming continuity of $u^{(iv)}$ and using the Intermediate Value Theorem, show that the *truncation error* is given by

$$\tau_i^h \equiv f(x_i) - (A^h\mathbf{u})_i = \frac{h^2}{24}(u^{(iv)}(\xi^+) + u^{(iv)}(\xi^-)) = \frac{h^2}{12}f''(\xi_i),$$

where $x_{i-1} < \xi_i < x_{i+1}$.

(c) Recall that if A is symmetric, then $\|A\|_2 = \rho(A)$, and that if $A\mathbf{w} = \lambda\mathbf{w}$, then $A^{-1}\mathbf{w} = \lambda^{-1}\mathbf{w}$. Use the eigenvalues of A^h and the definition of matrix norms to show that $\|(A^h)^{-1}\|_h \le \gamma$, where $\gamma \approx \pi^{-2}$. Note that this bound is independent of h.

(d) Show that if f'' is bounded on $[0,1]$ and $v_i = f''(\xi_i)$, then $\|\mathbf{v}\|_h$ is bounded. Is this result true for the (unscaled) Euclidean norm?

(e) Combine the above facts and use properties of matrix norms to conclude that the discrete L^2 norm of the discretization error is bounded by a constant times h^2; begin by making the observation that

$$\|E^h\|_h = \|\mathbf{u}^h - \mathbf{u}\|_h = \|(A^h)^{-1}A^h(\mathbf{u}^h - \mathbf{u})\|_h.$$

4. **Effect of full weighting.** Verify that $I_h^{2h}\mathbf{w}_k^h = \cos^2\left(\frac{k\pi}{2n}\right)\mathbf{w}_k^{2h}$, where $w_{k,j}^h = \sin\left(\frac{jk\pi}{n}\right)$, $1 \le k < \frac{n}{2}$, and I_h^{2h} is the full weighting operator.

5. **Effect of full weighting.** Verify that $I_h^{2h}\mathbf{w}_{k'}^h = -\sin^2\left(\frac{k\pi}{2n}\right)\mathbf{w}_k^{2h}$, where $k' = n-k$, $1 \le k < \frac{n}{2}$, and I_h^{2h} is the full weighting operator. What happens to $\mathbf{w}_{n/2}^h$ under full weighting?

6. **Complementary modes.** Show that the complementary modes $\{\mathbf{w}_k^h, \mathbf{w}_{k'}^h\}$ on Ω^h are related by $w_{k',j}^h = (-1)^{j+1}w_{k,j}^h$, where $k' = n-k$.

7. **Null space of full weighting.** Verify that the vectors $\mathbf{n}_j = A^h\hat{\mathbf{e}}_j^h$, where j is odd and $\hat{\mathbf{e}}_j^h$ is the jth unit vector on Ω^h, form a basis for the null space of I_h^{2h}.

8. **Effect of linear interpolation.** Show that $I_{2h}^h\mathbf{w}_k^{2h} = c_k\mathbf{w}_k^h - s_k\mathbf{w}_{k'}^h$, where $1 \le k < \frac{n}{2}$, $k' = n-k$, $c_k = \cos^2\left(\frac{k\pi}{2n}\right)$, $s_k = \sin^2\left(\frac{k\pi}{2n}\right)$, and I_{2h}^h is the linear interpolation operator.

9. **Two-grid correction scheme.** Using the previous results concerning the effects of A^h, I_h^{2h}, and I_{2h}^h on the fundamental modes, determine the effect of the coarse-grid correction operator without relaxation, $TG = I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h$, on the modes $\mathbf{w}_k$ and $\mathbf{w}_{k'}$, where $1 \le k \le \frac{n}{2}$, $k' = n-k$. (Recall that if $A\mathbf{w} = \lambda\mathbf{w}$, then $A^{-1}\mathbf{w} = \lambda^{-1}\mathbf{w}$.)

10. **Effect of two-grid correction.** Let R be an iteration matrix with the same eigenvectors, $\mathbf{w}_k$, as A^h and with eigenvalues λ_k. Consider the coarse-grid correction operator with ν sweeps of relaxation, $TG = [I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h] R^\nu$.

(a) Determine how TG acts on the modes $\mathbf{w}_k$ and $\mathbf{w}_{k'}$, where $1 \le k \le \frac{n}{2}$, $k' = n - k$.

(b) Find the specific convergence factors for the coarse-grid correction operator that uses one sweep of damped Jacobi with $\omega = \frac{2}{3}$.

11. **An important equivalence.** Prove that the null space of TG is the range of interpolation, where TG is the coarse grid correction operator without relaxation. Equation (5.12) of the text can be used to show that $N(TG) \supset \mathcal{R}(I_{2h}^h)$. Use equation (5.13) and the dimensions of the subspaces to show that $N(TG) = \mathcal{R}(I_{2h}^h)$.

12. **Two-dimensional five-point stencil.** When the two-dimensional model problem is discretized on a uniform grid with $h_x = h_y = h$, the coefficients at each grid point are given by the five-point stencil

$$A^h = \frac{1}{h^2} \begin{pmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{pmatrix}.$$

What does the stencil for $A^{2h} = I_h^{2h} A^h I_{2h}^h$ look like if I_{2h}^h is based on bilinear interpolation and I_h^{2h} is based on (a) full weighting? (b) injection?

13. **Two-dimensional nine-point stencil.** Repeat the previous problem with the nine-point stencil

$$A^h = \frac{1}{3h^2} \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}.$$