Chapter 9

Multigrid Methods

9.1 The residual equation

Definition 9.1. In solving a linear system $A\mathbf{x} = \mathbf{b}$, the *error of an approximate solution* $\tilde{\mathbf{x}}$ is

$$\mathbf{e}(\tilde{\mathbf{x}}) := \mathbf{x} - \tilde{\mathbf{x}} \tag{9.1}$$

and the *residual* of $\tilde{\mathbf{x}}$ is

$$\mathbf{r}(\tilde{\mathbf{x}}) := \mathbf{b} - A\tilde{\mathbf{x}}. \tag{9.2}$$

Lemma 9.2. The error and the residual of an approximate solution $\tilde{\mathbf{x}}$ satisfy the *residual equation*

$$A\mathbf{e} = \mathbf{r}. \tag{9.3}$$

Proof. This follows from Definition 9.1 in the same way that Lemma 7.17 follows from Lemma 7.16. □

Definition 9.3. The *condition number of a matrix* A is

$$\operatorname{cond}(A) := \|A\|_2 \|A^{-1}\|_2. \tag{9.4}$$

Theorem 9.4. The relative error of an approximate solution is bounded by its relative residual.

$$\frac{1}{\operatorname{cond}(A)} \frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2} \le \frac{\|\mathbf{e}\|_2}{\|\mathbf{x}\|_2} \le \operatorname{cond}(A) \frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2}. \tag{9.5}$$

Exercise 9.5. Prove Theorem 9.4.

9.2 The model problem

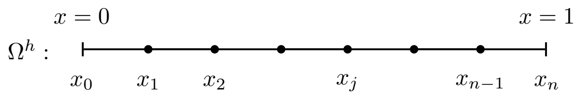
Definition 9.6. The *model problem* for our exposition of multigrid methods is the one-dimensional Poisson equation with homogeneous boundary condition

$$\begin{cases} -\Delta u = f & \text{in } \Omega := (0,1); \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{9.6}$$

Example 9.7. As a special case of Example 7.9 with $\alpha = 0$, $\beta = 0$, and $m = n - 1$, our discretization of (9.6) yields a linear system

$$A\mathbf{u} = \mathbf{f}, \tag{9.7}$$

where the unit interval Ω is discretized by uniform grid size $h = \frac{1}{n}$ into n cells with cell boundaries at $x_j = jh = \frac{j}{n}$ for $j = 0, 1, \ldots, n$, the knowns $f_j = f(x_j)$ and the unknowns u_j are located at the internal nodes x_j with $j = 1, 2, \ldots, n-1$.



The matrix $A \in \mathbb{R}^{(n-1)\times(n-1)}$ is the same as that in (7.11), i.e., a Toeplitz matrix given by

$$a_{ij} = \begin{cases} \frac{2}{h^2} & \text{if } i = j; \\ -\frac{1}{h^2} & \text{if } i - j = \pm 1; \\ 0, & \text{otherwise.} \end{cases} \tag{9.8}$$

By Lemma 7.24, the eigenvalues and eigenvectors of A are

$$\lambda_k(A) = \frac{4}{h^2}\sin^2\frac{k\pi}{2n} = \frac{4}{h^2}\sin^2\frac{kh\pi}{2}, \tag{9.9}$$

$$w_{k,j} = \sin\frac{jk\pi}{n} = \sin(x_j k\pi), \tag{9.10}$$

where $j, k = 1, 2, \ldots, n-1$.

Exercise 9.8. What are the values of $\operatorname{cond}(A)$ for A in (7.11) for $n = 8$ and $n = 1024$?

9.3 Algorithmic components

9.3.1 Fourier modes on Ω^h

Notation 7. Ω^h denotes the uniform grid of n intervals that discretizes the problem domain Ω. Occasionally we also abuse the notation to mean the corresponding vector space of grid functions $\{\Omega^h \to \mathbb{R}\}$.

Definition 9.9. The *wavelength of a sinusoidal function* is the distance of one sinusoidal period. The *wavenumber of a sinusoidal function* k is the number of half sinusoidal waves in unit length.

Lemma 9.10. The kth Fourier mode with its jth component as $w_{k,j} = \sin(x_j k\pi)$ has wavelength $L = \frac{2}{k}$.

Proof. By Definition 9.9, $\sin(x_j k\pi) = -\sin(x_j + \frac{L}{2})k\pi$ implies $x_j k\pi = (x_j + \frac{L}{2})k\pi - \pi$. Hence $k = \frac{2}{L}$. □

Exercise 9.11. For $\Omega = (0, 1)$, plot to show that the maximum wavenumber that is representable on Ω^h is $n_{\max} = \frac{1}{h}$. What if we require that the Fourier mode be 0 at the boundary points?

Proof. Alternate from local maximum and local minimum at the $n+1$ grid points and we have $n_{\max} = \frac{1}{h}$. When homogeneous Dirichlet conditions are imposed on the boundary points, the maximum number of alternation between local extrema is reduced by one. Hence we have $n_{\max} = \frac{1}{h} - 1$. □

Lemma 9.12 (Aliasing)**.** For $k \in (n, 2n)$ on Ω^h, the Fourier mode $\mathbf{w}_k$ of which the jth component is $w_{k,j} = \sin(x_j k\pi)$ is actually represented as the additive inverse of the mode $\mathbf{w}_{k'}$ where $k' = 2n - k$.

Proof. It is readily verified that

$$\begin{aligned} \sin(x_j k\pi) &= -\sin(2j\pi - x_j k\pi) = -\sin(x_j(2n-k)\pi) \\ &= -\sin(x_j k'\pi) = -w_{k',j}. \end{aligned}$$

□

Example 9.13. According to Lemma 9.12, the mode with $k = \frac{3}{2}n$ is represented by $k = \frac{1}{2}n$.

Exercise 9.14. Plot the case of $n = 6$ for Example 9.13.

Definition 9.15. On Ω^h, the Fourier modes with wavenumbers $k \in [1, \frac{n}{2})$ are called the *low-frequency* (LF) or *smooth* modes, those with $k \in [\frac{n}{2}, n)$ the *high-frequency* (HF) or *oscillatory* modes.

9.3.2 Relaxation

Lemma 9.16. For the linear system (9.7), the weighted Jacobi in Definition 8.9 has the iteration matrix

$$T_\omega = (1-\omega)I + \omega D^{-1}(L+U) = I - \frac{\omega h^2}{2}A, \tag{9.11}$$

whose eigenvectors are the same as those of A, with the corresponding eigenvalues as

$$\lambda_k(T_\omega) = 1 - 2\omega \sin^2 \frac{k\pi}{2n}, \tag{9.12}$$

where $k = 1, 2, \ldots, n-1$.

Exercise 9.17. Prove Lemma 9.16.

Exercise 9.18. Write a program to reproduce Fig. 2.7 in the book by Briggs et al. [2000]. For $n = 64$, $\omega \in [0, 1]$, verify $\rho(T_\omega) \geq 0.9986$ and hence slow convergence.

Definition 9.19. The *smoothing factor* μ is the maximal factor of damping for HF modes. An iterative method is said to have the *smoothing property* if μ is small and independent of the grid size.

Example 9.20. The smoothing factor of the weighted Jacobi is determined by the optimization problem

$$\mu = \min_{\omega \in (0,1]} \max_{k \in [\frac{n}{2}, n)} |\lambda_k(T_\omega)|. \tag{9.13}$$

Since $\lambda_k(T_\omega)$ is a monotonically decreasing function, the minimum is obtained by setting $\lambda_{\frac{n}{2}}(T_\omega) = -\lambda_n(T_\omega)$, which implies $\omega = \frac{2}{3}$. Consequently we have $|\lambda_k| \leq \mu = \frac{1}{3}$.

Exercise 9.21. Write a program to reproduce Figure 2.8 in the book by Briggs et al. [2000], verifying that regular Jacobi is only good for damping modes $16 \leq k \leq 48$. In contrast, for $\omega = \frac{2}{3}$, the modes $16 \leq k < 64$ are all damped out quickly.

9.3.3 Restriction and prolongation

Lemma 9.22. The kth LF mode on Ω^h becomes the kth mode (LF or HF) on Ω^{2h}:

$$w^h_{k,2j} = w^{2h}_{k,j}. \tag{9.14}$$

LF modes $k \in [\frac{n}{4}, \frac{n}{2})$ of Ω^h will become HF modes on Ω^{2h}.

Proof. It is readily verified that

$$w^h_{k,2j} = \sin\frac{2jk\pi}{n} = \sin\frac{jk\pi}{\frac{n}{2}} = w^{2h}_{k,j}, \tag{9.15}$$

where $k \in [1, \frac{n}{2})$. Because of the smaller range of k on Ω^{2h}, the modes with $k \in [\frac{n}{4}, \frac{n}{2})$ are HF by definition since the highest wavenumber is $\frac{n}{2}$ on Ω^{2h}. □

Definition 9.23. The *restriction* operator

$$I_h^{2h} : \mathbb{R}^{n-1} \to \mathbb{R}^{\frac{n}{2}-1}$$

maps a vector on the fine grid Ω^h to its counterpart on the coarse grid Ω^{2h}:

$$I_h^{2h}\mathbf{v}^h = \mathbf{v}^{2h}. \tag{9.16}$$

Definition 9.24. The *injection* operator is a restriction operator given by

$$v_j^{2h} = v_{2j}^h, \tag{9.17}$$

where $j = 1, 2, \ldots, \frac{n}{2} - 1$.

Definition 9.25. The *full-weighting* operator is a restriction operator given by

$$v_j^{2h} = \frac{1}{4}\left(v^h_{2j-1} + 2v^h_{2j} + v^h_{2j+1}\right), \tag{9.18}$$

where $j = 1, 2, \ldots, \frac{n}{2} - 1$.

Example 9.26. For $n = 8$, the full-weighting operator is

$$I_h^{2h} = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & \\ & & & & 1 & 2 & 1 \end{bmatrix}. \tag{9.19}$$

Definition 9.27. The *prolongation or interpolation* operator

$$I_{2h}^h : \mathbb{R}^{\frac{n}{2}-1} \to \mathbb{R}^{n-1}$$

maps a vector on the coarse grid Ω^{2h} to its counterpart on the fine grid Ω^h:

$$I_{2h}^h\mathbf{v}^{2h} = \mathbf{v}^h. \tag{9.20}$$

Definition 9.28. The *linear interpolation* operator is a prolongation operator given by

$$\begin{aligned} v^h_{2j} &= v^{2h}_j, \\ v^h_{2j+1} &= \tfrac{1}{2}(v^{2h}_j + v^{2h}_{j+1}). \end{aligned} \tag{9.21}$$

Example 9.29. For $n = 8$, the linear interpolation operator is

$$I_{2h}^{h} = \frac{1}{2}\begin{bmatrix} 1 & & \\ 2 & & \\ 1 & 1 & \\ & 2 & \\ & 1 & 1 \\ & & 2 \\ & & 1 \end{bmatrix}. \tag{9.22}$$

9.3.4 Two-grid correction

Definition 9.30. The *two-grid correction scheme*

$$\mathbf{v}^h \leftarrow \texttt{TG}(\mathbf{v}^h, \mathbf{f}^h, \nu_1, \nu_2) \tag{9.23}$$

solves $A\mathbf{u} = \mathbf{f}$ in (9.7) via steps as follows.

(TG-1) Relax $A^h\mathbf{u}^h = \mathbf{f}^h$ for ν_1 times on Ω^h with initial guess $\mathbf{v}^h$: $\mathbf{v}^h \leftarrow T_\omega^{\nu_1}\mathbf{v}^h + \mathbf{c}'(f)$,

(TG-2) Compute the fine-grid residual $\mathbf{r}^h = \mathbf{f}^h - A^h\mathbf{v}^h$ and restrict it to the coarse grid by $\mathbf{r}^{2h} = I_h^{2h}\mathbf{r}^h$: $\mathbf{r}^{2h} \leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h)$,

(TG-3) Solve $A^{2h}\mathbf{e}^{2h} = \mathbf{r}^{2h}$ on Ω^{2h}: $\mathbf{e}^{2h} \leftarrow (A^{2h})^{-1}\mathbf{r}^{2h}$,

(TG-4) 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: $\mathbf{v}^h \leftarrow \mathbf{v}^h + I_{2h}^h\mathbf{e}^{2h}$,

(TG-5) Relax $A^h\mathbf{u}^h = f^h$ for ν_2 times on Ω^h with initial guess $\mathbf{v}^h$: $\mathbf{v}^h \leftarrow T_\omega^{\nu_2}\mathbf{v}^h + \mathbf{c}'(f)$.

Lemma 9.31. Acting on the error vector, the iteration matrix of the two-grid correction scheme (9.23) is

$$TG = T_\omega^{\nu_2}\left[I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\right]T_\omega^{\nu_1}. \tag{9.24}$$

Proof. By Definition 9.30, the residual on the fine grid is

$$\mathbf{r}^h(\mathbf{v}^h) = \mathbf{f}^h - A^h\left(T_\omega^{\nu_1}\mathbf{v}^h + \mathbf{c}'(f)\right).$$

The two-grid correction scheme with $\nu_2 = 0$ replaces the initial guess with

$$\mathbf{v}^h \leftarrow T_\omega^{\nu_1}\mathbf{v}^h + \mathbf{c}'(f) + I_{2h}^h(A^{2h})^{-1}I_h^{2h}\mathbf{r}^h(\mathbf{v}^h)$$

which also holds for the exact solution $\mathbf{u}^h$

$$\mathbf{u}^h \leftarrow T_\omega^{\nu_1}\mathbf{u}^h + \mathbf{c}'(f) + I_{2h}^h(A^{2h})^{-1}I_h^{2h}\mathbf{r}^h(\mathbf{u}^h).$$

Subtracting the two equations yields

$$\mathbf{e}^h \leftarrow T_\omega^{\nu_1}\mathbf{e}^h - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^hT_\omega^{\nu_1}\mathbf{e}^h.$$

Similar arguments applied to step (TG-5) yield (9.24). □

9.3.5 Multigrid cycles

Definition 9.32. The *V-cycle scheme*

$$\mathbf{v}^h \leftarrow \texttt{VC}^h(\mathbf{v}^h, \mathbf{f}^h, \nu_1, \nu_2) \tag{9.25}$$

solves $A\mathbf{u} = \mathbf{f}$ in (9.7) via steps as follows.

(VC-1) Relax ν_1 times on $A^h\mathbf{u}^h = \mathbf{f}^h$ with initial guess $\mathbf{v}^h$.

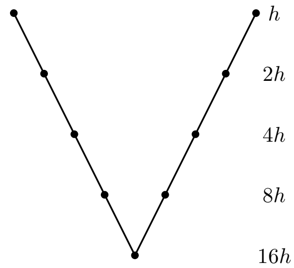
(VC-2) If Ω^h is the coarsest grid, go to (VC-4), otherwise

$$\begin{aligned} \mathbf{f}^{2h} &\leftarrow I_h^{2h}(\mathbf{f}^h - A^h\mathbf{v}^h), \\ \mathbf{v}^{2h} &\leftarrow \mathbf{0}, \\ \mathbf{v}^{2h} &\leftarrow \texttt{VC}^{2h}(\mathbf{v}^{2h}, \mathbf{f}^{2h}, \nu_1, \nu_2). \end{aligned}$$

(VC-3) Interpolate error back and correct the solution:

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

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



Lemma 9.33. In a D-dimensional domain with $n = 2^m$ cells ($m \in \mathbb{N}^+$) along each dimension, the storage cost of V-cycles is

$$2n^{\mathrm{D}}\left(1 + 2^{-\mathrm{D}} + 2^{-2\mathrm{D}} + \cdots + 2^{-m\mathrm{D}}\right) < \frac{2n^{\mathrm{D}}}{1 - 2^{-\mathrm{D}}}. \tag{9.26}$$

Let WU denote the computational cost of performing one relaxation sweep on the finest grid. After neglecting the intergrid transfer, the computational cost of a single V-cycle with $\nu_1 = \nu_2 = 1$ is

$$2\mathrm{WU}\left(1 + 2^{-\mathrm{D}} + 2^{-2\mathrm{D}} + \cdots + 2^{-m\mathrm{D}}\right) < \frac{2}{1 - 2^{-\mathrm{D}}}\mathrm{WU}. \tag{9.27}$$

Proof. On each grid, both vectors of errors and residuals must be stored, and this justifies the factor of 2 in (9.26); the rest of (9.26) follows from Definition 9.32. A similar argument yields (9.27). □

Definition 9.34. The *full multigrid V-cycle*

$$\mathbf{v}^h \leftarrow \texttt{FMG}^h(\mathbf{f}^h, \nu_1, \nu_2) \tag{9.28}$$

solves $A\mathbf{u} = \mathbf{f}$ in (9.7) via steps as follows.

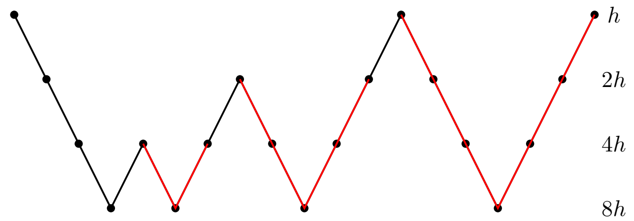
(FMG-1) If Ω^h is the coarsest grid, set $\mathbf{v}^h \leftarrow \mathbf{0}$ and go to (FMG-3), otherwise

$$\begin{aligned} \mathbf{f}^{2h} &\leftarrow I_h^{2h}\mathbf{f}^{\mathbf{h}}, \\ \mathbf{v}^{2h} &\leftarrow \texttt{FMG}^{2h}(\mathbf{f}^{2h}, \nu_1, \nu_2). \end{aligned}$$

(FMG-2) Correct $\mathbf{v}^h \leftarrow I_{2h}^h\mathbf{v}^{2h}$.

(FMG-3) Perform a V-cycle with the initial guess as $\mathbf{v}^h$:

$$\mathbf{v}^h \leftarrow \texttt{VC}^h(\mathbf{v}^h, \mathbf{f}^h, \nu_1, \nu_2).$$



Exercise 9.35. Show that, for $\nu_1 = \nu_2 = 1$, the computational cost of an FMG cycle is less than $\frac{2}{(1-2^{-D})^2}$ WU. Give upper bounds as tight as possible for computational costs of an FMG cycle for $D = 1, 2, 3$.

9.4 Convergence analysis

9.4.1 The spectral picture

Definition 9.36. A Fourier mode $\mathbf{w}_k^h$ with $k \in [1, \frac{n}{2})$ and the mode $\mathbf{w}_{k'}^h$ with $k' = n - k$ are called *complementary modes* on Ω^h.

Lemma 9.37. A pair of complementary modes satisfy

$$w_{k',j}^h = (-1)^{j+1} w_{k,j}^h. \tag{9.29}$$

Proof. This follows from

$$w_{k',j}^h = \sin\frac{(n-k)j\pi}{n} = \sin\left(j\pi - \frac{kj\pi}{n}\right) = (-1)^{j+1} w_{k,j}^h.$$

□

Lemma 9.38. The action of the full-weighting operator on a pair of complementary modes on Ω^h is

$$I_h^{2h}\mathbf{w}_k^h = c_k \mathbf{w}_k^{2h} := \cos^2\frac{k\pi}{2n}\mathbf{w}_k^{2h}, \tag{9.30a}$$

$$I_h^{2h}\mathbf{w}_{k'}^h = -s_k \mathbf{w}_k^{2h} := -\sin^2\frac{k\pi}{2n}\mathbf{w}_k^{2h}, \tag{9.30b}$$

where $k \in [1, \frac{n}{2})$, $k' = n - k$. In addition, $I_h^{2h}\mathbf{w}_{\frac{n}{2}}^h = \mathbf{0}$.

Proof. For the smooth mode k, we have

$$\begin{aligned}
&\left(I_h^{2h}\mathbf{w}_k^h\right)_j \\
=&\frac{1}{4}\sin\frac{(2j-1)k\pi}{n} + \frac{1}{2}\sin\frac{2jk\pi}{n} + \frac{1}{4}\sin\frac{(2j+1)k\pi}{n} \\
=&\frac{1}{2}\left(1+\cos\frac{k\pi}{n}\right)\sin\frac{2jk\pi}{n} = \cos^2\frac{k\pi}{2n} w_{k,j}^{2h},
\end{aligned}$$

where the last step follows from Lemma 9.22. (9.30b) can be proved by similar steps by replacing k with $n-k$. □

Lemma 9.39. The action of the linear interpolation operator on Ω^{2h} is

$$I_{2h}^h \mathbf{w}_k^{2h} = c_k \mathbf{w}_k^h - s_k \mathbf{w}_{k'}^h, \tag{9.31}$$

where $k' = n - k$.

Proof. Lemma 9.37 and trigonometric identities yield

$$\begin{aligned}
c_k w_{k,j}^h - s_k w_{k',j}^h &= \left(\cos^2\frac{k\pi}{2n} + (-1)^j \sin^2\frac{k\pi}{2n}\right) w_{k,j}^h \\
&= \begin{cases} w_{k,j}^h & \text{if } j \text{ is even;} \\ \cos\frac{k\pi}{n} w_{k,j}^h & \text{if } j \text{ is odd.} \end{cases}
\end{aligned}$$

On the other hand, by Definition 9.28, we have

$$\left(I_{2h}^h \mathbf{w}_k^{2h}\right)_j = \begin{cases} w_{k,j}^h, & \text{if } j \text{ is even,} \\ \frac{1}{2}\sin\frac{k\pi(j-1)}{n} + \frac{1}{2}\sin\frac{k\pi(j+1)}{n} & \text{if } j \text{ is odd,} \end{cases}$$

where last expression simplifies to $\cos\frac{k\pi}{n} w_{k,j}^h$. □

Theorem 9.40. The two-grid correction operator is invariant on the subspace $W_k^h = \mathrm{span}\{\mathbf{w}_k^h, \mathbf{w}_{k'}^h\}$.

$$TG\mathbf{w}_k = \lambda_k^{\nu_1+\nu_2} s_k \mathbf{w}_k + \lambda_k^{\nu_1}\lambda_{k'}^{\nu_2} s_k \mathbf{w}_{k'} \tag{9.32a}$$

$$TG\mathbf{w}_{k'} = \lambda_{k'}^{\nu_1}\lambda_k^{\nu_2} c_k \mathbf{w}_k + \lambda_{k'}^{\nu_1+\nu_2} c_k \mathbf{w}_{k'}, \tag{9.32b}$$

where λ_k is the eigenvalue of T_ω.

Proof. Recall from (9.9) that $\frac{4}{h^2}s_k$ is the eigenvalue of A^h. Consider first the case of $\nu_1 = \nu_2 = 0$.

$$A^h\mathbf{w}_k^h = \frac{4s_k}{h^2}\mathbf{w}_k^h \tag{9.33a}$$

$$\Rightarrow I_h^{2h}A^h\mathbf{w}_k^h = \frac{4c_k s_k}{h^2}\mathbf{w}_k^{2h} \tag{9.33b}$$

$$\Rightarrow (A^{2h})^{-1}I_h^{2h}A^h\mathbf{w}_k^h = \frac{4c_k s_k}{h^2}\frac{(2h)^2}{4\sin^2\frac{k\pi}{n}}\mathbf{w}_k^{2h} = \mathbf{w}_k^{2h} \tag{9.33c}$$

$$\Rightarrow -I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\mathbf{w}_k^h = -c_k\mathbf{w}_k^h + s_k\mathbf{w}_{k'}^h \tag{9.33d}$$

$$\Rightarrow \left[I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\right]\mathbf{w}_k^h = s_k\mathbf{w}_k^h + s_k\mathbf{w}_{k'}^h. \tag{9.33e}$$

Similarly, we have

$$A^h\mathbf{w}_{k'}^h = \frac{4s_{k'}}{h^2}\mathbf{w}_{k'}^h = \frac{4c_k}{h^2}\mathbf{w}_{k'}^h \tag{9.34a}$$

$$\Rightarrow I_h^{2h}A^h\mathbf{w}_{k'}^h = -\frac{4c_k s_k}{h^2}\mathbf{w}_k^{2h} \tag{9.34b}$$

$$\Rightarrow (A^{2h})^{-1}I_h^{2h}A^h\mathbf{w}_{k'}^h = -\frac{4c_k s_k}{h^2}\frac{(2h)^2}{4\sin^2\frac{k\pi}{n}}\mathbf{w}_k^{2h} = -\mathbf{w}_k^{2h} \tag{9.34c}$$

$$\Rightarrow -I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\mathbf{w}_{k'}^h = c_k\mathbf{w}_k^h - s_k\mathbf{w}_{k'}^h \tag{9.34d}$$

$$\Rightarrow \left(I - I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^h\right)\mathbf{w}_{k'}^h = c_k\mathbf{w}_k^h + c_k\mathbf{w}_{k'}^h, \tag{9.34e}$$

where $c_k = s_{k'}$ is applied in (9.34a).

Adding pre-smoothing incurs a scaling of $\lambda_k^{\nu_1}$ for (9.33e) and $\lambda_{k'}^{\nu_1}$ for (9.34e). In contrast, adding post-smoothing incurs a scaling of $\lambda_k^{\nu_2}$ for $\mathbf{w}_k^h$ and a scaling of $\lambda_{k'}^{\nu_2}$ for $\mathbf{w}_{k'}^h$ in both (9.33e) and (9.34e). Hence (9.32) holds. □

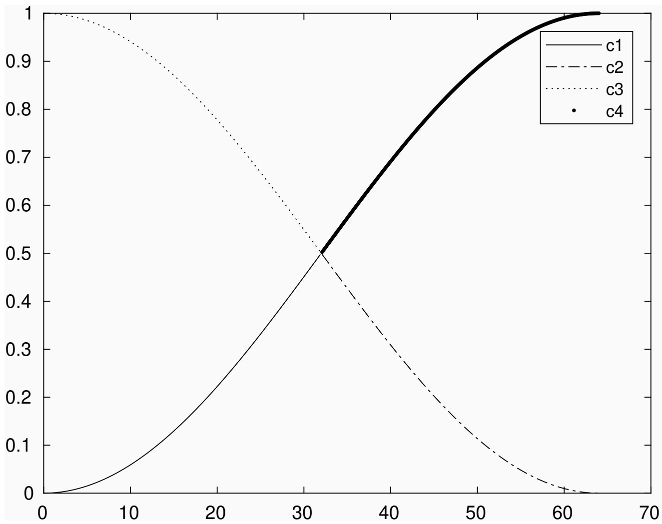
Exercise 9.41. Rewrite (9.32) as

$$TG\begin{bmatrix}\mathbf{w}_k \\ \mathbf{w}_{k'}\end{bmatrix} = \begin{bmatrix}\lambda_k^{\nu_1+\nu_2}s_k & \lambda_k^{\nu_1}\lambda_{k'}^{\nu_2}s_k \\ \lambda_{k'}^{\nu_1}\lambda_k^{\nu_2}c_k & \lambda_{k'}^{\nu_1+\nu_2}c_k\end{bmatrix}\begin{bmatrix}\mathbf{w}_k \\ \mathbf{w}_{k'}\end{bmatrix} = \begin{bmatrix}c_1 & c_2 \\ c_3 & c_4\end{bmatrix}\begin{bmatrix}\mathbf{w}_k \\ \mathbf{w}_{k'}\end{bmatrix}. \tag{9.35}$$

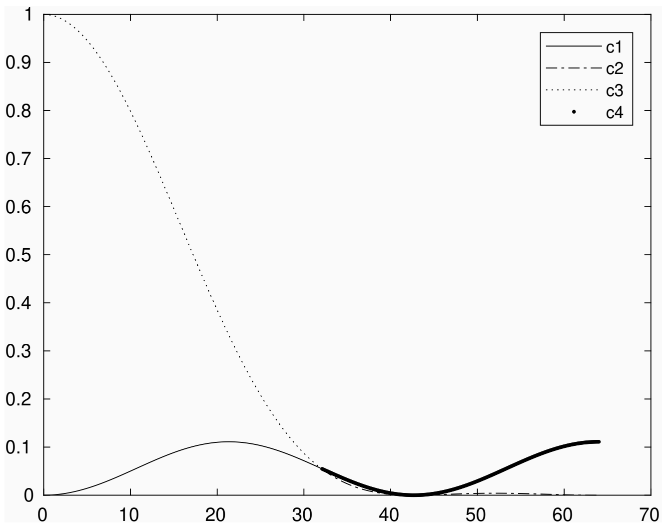
Explain why the magnitude of all four c_i's are small. Deduce the main conclusion $\rho(TG) \approx 0.1$ by reproducing the following plots of the damping coefficients of two-grid correction with weighted Jacobi for $n = 64$ and $\omega = \frac{2}{3}$. The x-axis represents the wave number k. Repeat the plots for $n = 128$ to show the independence of $\rho(TG) \approx 0.1$ from the grid size.

Hint: It is tricky to plot the coefficients defined in (9.35). Since c_2, c_4 act on HF modes, one has to ensure that the components in the vectors s_k and c_k indeed correspond to those in $\mathbf{w}_{k'}$. If s_k and c_k are computed from an increasing order of the frequencies, then their components will have to be reversed for plotting. Physical intuition helps in this case: c_1 and c_4 should form one curve while c_2 and c_3 should form another.

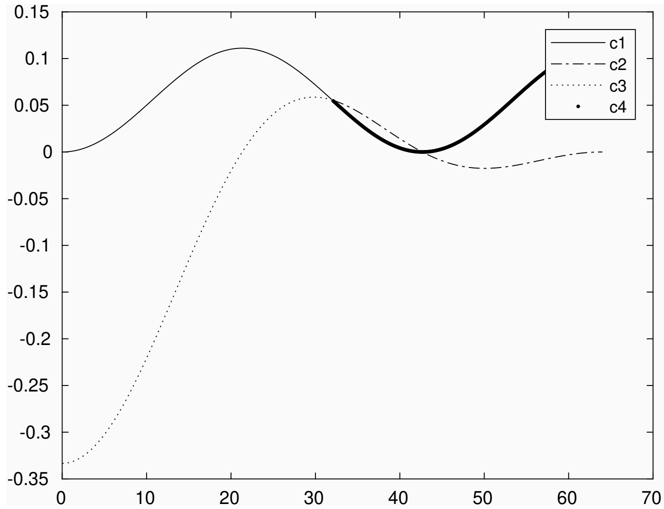
$\nu_1 = 0$, $\nu_2 = 0$.



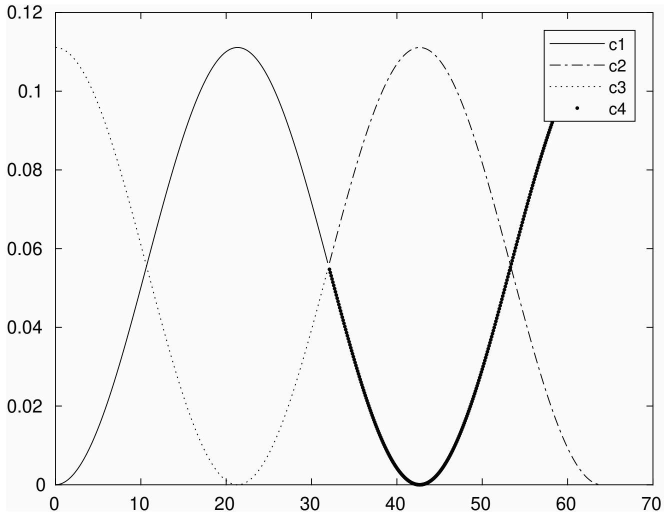
$\nu_1 = 0$, $\nu_2 = 2$.



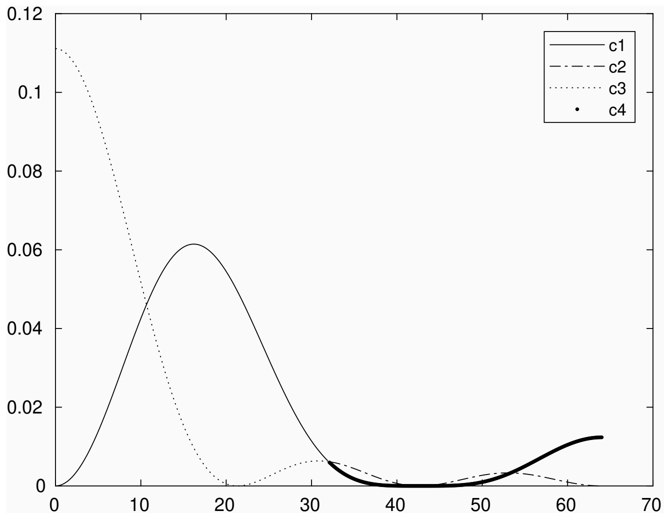
$\nu_1 = 1$, $\nu_2 = 1$.



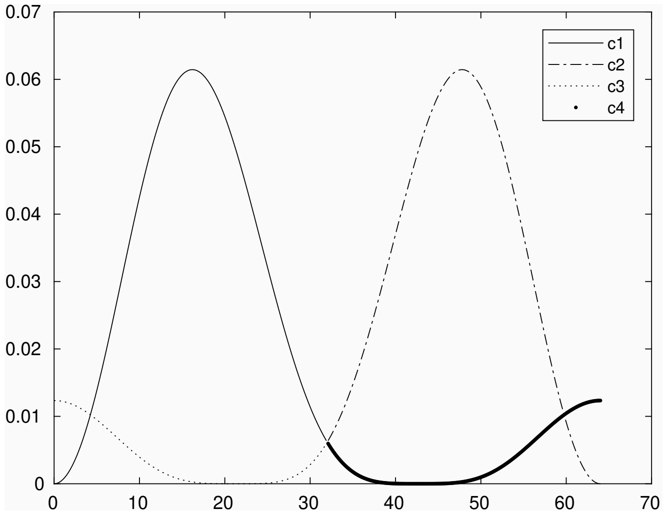
$\nu_1 = 2$, $\nu_2 = 0$.



$\nu_1 = 2$, $\nu_2 = 2$.



$\nu_1 = 4$, $\nu_2 = 0$.



9.4.2 The algebraic picture

Lemma 9.42. The full-weighting operator and the linear interpolation operator satisfy the *variational properties*

$$I_{2h}^h = c(I_h^{2h})^T, \tag{9.36a}$$

$$I_h^{2h} A^h I_{2h}^h = A^{2h}, \tag{9.36b}$$

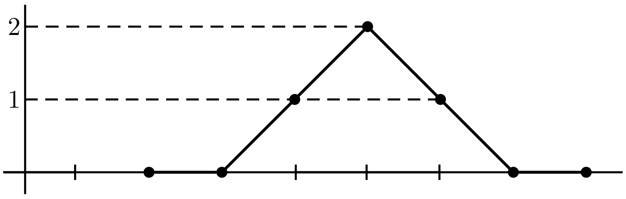
where $c = 2$ and the property (9.36b) is also called the *Galerkin condition.*

Proof. The conclusions follow from (9.8) and Definitions 9.25 and 9.28. □

Lemma 9.43. A basis for the range of the linear interpolation operator $\mathcal{R}(I_{2h}^h)$ is given by its columns, hence the range and the null space of I_{2h}^h satisfy

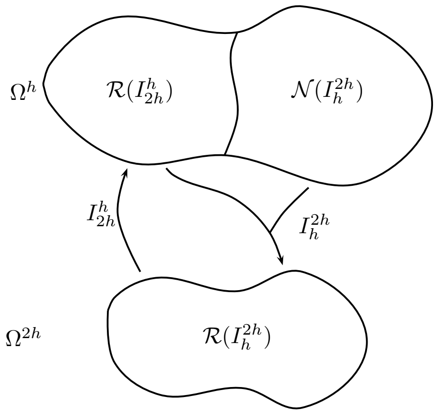
$$\dim \mathcal{R}\left(I_{2h}^h\right) = \frac{n}{2} - 1, \quad \mathcal{N}\left(I_{2h}^h\right) = \{\mathbf{0}\}. \tag{9.37}$$

Proof. $\mathcal{R}(I_{2h}^h) = \{I_{2h}^h \mathbf{v}^{2h} : \mathbf{v}^{2h} \in \Omega^{2h}\}$. The maximum dimension of $\mathcal{R}(I_{2h}^h)$ is thus $\frac{n}{2} - 1$. Any $\mathbf{v}^{2h}$ can be expressed as $\mathbf{v}^{2h} = \sum v_j^{2h} \mathbf{e}_j^{2h}$. It is obvious that the columns of I_{2h}^h are linearly independent. □



Lemma 9.44. The full-weighting operator satisfies

$$\dim \mathcal{R}(I_h^{2h}) = \frac{n}{2} - 1, \qquad \dim \mathcal{N}(I_h^{2h}) = \frac{n}{2}. \tag{9.38}$$



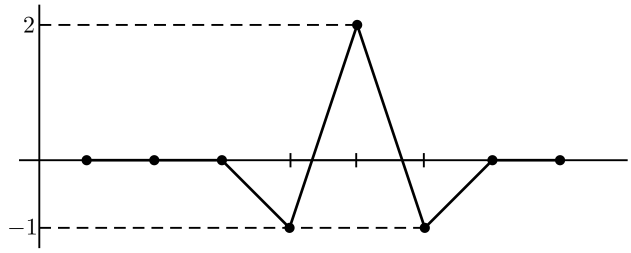
Exercise 9.45. Prove Lemma 9.44.

Lemma 9.46. A basis for the null space of the full-weighting operator is given by

$$\mathcal{N}(I_h^{2h}) = \mathrm{span}\{A^h \mathbf{e}_j^h : j \text{ is odd}\}, \tag{9.39}$$

where $\mathbf{e}_j^h$ is the jth unit vector on Ω^h.

Proof. Consider $I_h^{2h} A^h$. The jth row of I_h^{2h} has $2(j-1)$ leading zeros and the next three nonzero entries are $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$. Since the bandwidth of A^h is 3, it suffices to consider only five columns of A^h for potentially non-zero dot-product $\sum_i (I_h^{2h})_{ji}(A^h)_{ik}$. For $2j \pm 1$, these dot products are zero; for $2j$, the dot product is $\frac{1}{2}$; for $2j \pm 2$, the dot product is $-\frac{1}{4}$. Hence for any odd j, we have $I_h^{2h} A^h \mathbf{e}_j^h = \mathbf{0}$. □



Theorem 9.47. The null space of the two-grid correction operator (without relaxation) is the range of linear interpolation:

$$\mathcal{N}(TG) = \mathcal{R}(I_{2h}^h). \tag{9.40}$$

Proof. If $\mathbf{s}^h \in \mathcal{R}(I_{2h}^h)$, then $\mathbf{s}^h = I_{2h}^h \mathbf{q}^{2h}$.

$$TG\mathbf{s}^h = \left[I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h\right] I_{2h}^h \mathbf{q}^{2h} = \mathbf{0},$$

where the last step comes from (9.36b). Hence we have $\mathcal{R}(I_{2h}^h) \subseteq \mathcal{N}(TG)$. By Lemma 9.46, $\mathbf{t}^h \in \mathcal{N}(I_h^{2h} A^h)$ implies $\mathbf{t}^h = \sum_{j \text{ is odd}} t_j \mathbf{e}_j$. Consequently, we have

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

i.e., TG is the identity operator when acting on $\mathcal{N}(I_h^{2h} A^h)$. As shown in the plot below Lemma 9.44, the dimension of $\mathcal{N}(TG)$ is no greater than the dimension of $\mathcal{R}(I_h^{2h} A^h)$, which is the same as $\dim \mathcal{R}(I_{2h}^h)$ since A^h is a bijection with full rank on $\mathbb{R}^{n-1}$. This implies that $\dim\mathcal{N}(TG) \leq \dim\mathcal{R}(I_{2h}^h)$, which completes the proof. □

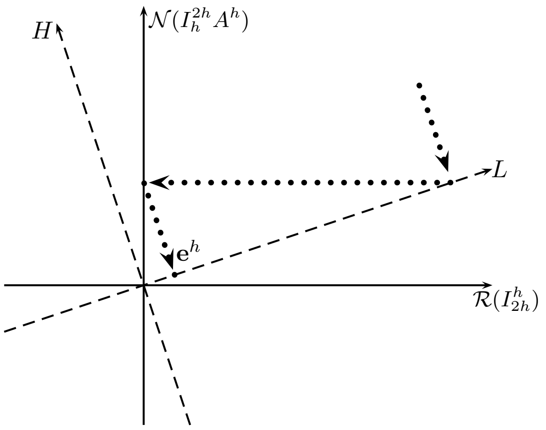
Definition 9.48. Let A be an $n\times n$ symmetric positive definite matrix. The *A-inner product* or *energy inner product* of two vectors $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ is defined as

$$\langle \mathbf{u},\mathbf{v}\rangle_A := \langle A\mathbf{u},\mathbf{v}\rangle, \tag{9.41}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product on $\mathbb{R}^n$. Naturally, the *A-norm* or *energy norm* is defined as

$$\|\mathbf{u}\|_A := \sqrt{\langle \mathbf{u},\mathbf{u}\rangle_A}. \tag{9.42}$$

Two vectors $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ are *A-orthogonal* iff $\langle \mathbf{u},\mathbf{v}\rangle_A = 0$.



9.4.3 The optimal complexity of FMG

Definition 9.49. The error of a computed value v_i^h from the corresponding exact solution $u(x_i)$ is

$$E_i^h := v_i^h - u(x_i) = v_i^h - u_i^h + u_i^h - u(x_i).$$

The *discretization error* is the error $u_i^h - u(x_i)$ incurred by truncating the Taylor series of exact values and the *algebraic error* is the error $v_i^h - u_i^h$ incurred by inexact solution of the linear system.

Lemma 9.50. When interpolating errors from a coarse grid to the fine grid, we have

$$c\left\|\mathbf{v}^{2h}-\mathbf{u}^{2h}\right\|_{A^{2h}} = \left\|I_{2h}^h\mathbf{v}^{2h} - I_{2h}^h\mathbf{u}^{2h}\right\|_{A^h}. \tag{9.43}$$

where $c\in\mathbb{R}^+$ is the constant in (9.36).

Proof. Definition 9.48 and Lemma 9.42 yield

$$\begin{aligned}
&\|\mathbf{v}^{2h}-\mathbf{u}^{2h}\|^2_{A^{2h}}\\
=&\left\langle A^{2h}(\mathbf{v}^{2h}-\mathbf{u}^{2h}),\mathbf{v}^{2h}-\mathbf{u}^{2h}\right\rangle\\
=&\left\langle I_h^{2h}A^hI_{2h}^h(\mathbf{v}^{2h}-\mathbf{u}^{2h}),\mathbf{v}^{2h}-\mathbf{u}^{2h}\right\rangle\\
=&\left\langle A^hI_{2h}^h(\mathbf{v}^{2h}-\mathbf{u}^{2h}),\frac{1}{c}I_{2h}^h(\mathbf{v}^{2h}-\mathbf{u}^{2h})\right\rangle\\
=&\frac{1}{c}\|I_{2h}^h\mathbf{v}^{2h}-I_{2h}^h\mathbf{u}^{2h}\|^2_{A^h},
\end{aligned}$$

where the second step and the third step follow from Lemma 9.42. ☐

Lemma 9.51. Suppose there exists a constant $K\in\mathbb{R}^+$ independent of the grid size h such that

$$\left\|I_{2h}^h\mathbf{u}^{2h}-\mathbf{u}^h\right\|_{A^h}\le Kh^p. \tag{9.44}$$

Then a single FMG cycle in Definition 9.34 reduces the algebraic error from $O(1)$ to $O(h^p)$, i.e.,

$$\left\|\mathbf{e}^h\right\|_{A^h}\le Kh^p, \tag{9.45}$$

where $p=2$ is the order of accuracy of the discrete Laplacian (9.8).

Proof. We prove (9.45) by induction. On the coarsest grid, FMG is exact and thus (9.45) holds for the induction basis.

For the induction hypothesis, we assume that the linear system on Ω^{2h} has been solved to the level of discretization error so that

$$(*):\quad \|\mathbf{e}^{2h}\|_{A^{2h}}\le K(2h)^p.$$

We need to show that $(*)$ implies (9.45).

The initial algebraic error on Ω^h is

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

which yields

$$\begin{aligned}
\left\|\mathbf{e}_0^h\right\|_{A^h} &\le \left\|I_{2h}^h\mathbf{v}^{2h}-I_{2h}^h\mathbf{u}^{2h}\right\|_{A^h}+\left\|I_{2h}^h\mathbf{u}^{2h}-\mathbf{u}^h\right\|_{A^h}\\
&= c\left\|\mathbf{v}^{2h}-\mathbf{u}^{2h}\right\|_{A^{2h}}+\left\|I_{2h}^h\mathbf{u}^{2h}-\mathbf{u}^h\right\|_{A^h}\\
&\le cK(2h)^p+Kh^p=(1+c2^p)Kh^p,
\end{aligned}$$

where the second step follows from Lemma 9.50 and the third step from (9.44) and the induction hypothesis. Then we have $1+c2^p=9$ from $p=2$ and $c=2$. Exercise 9.41 states that $\rho(TG)\approx 0.1$ for a $\mathtt{VC}(2,1)$-cycle and hence one V-cycle is enough to reduce $\left\|\mathbf{e}_0^h\right\|_{A^h}$ to less than Kh^p. ☐

Theorem 9.52. For the FD discretization (in Example 9.7) of the model problem in Definition 9.6, a single FMG cycle is sufficient to achieve second-order accuracy, with each computated result on Ω^h produced in $O(\frac{1}{h})$ time.

Proof. By Definition 9.49, we have

$$\|\mathbf{E}^h\|\le\|\mathbf{u}^h-\mathbf{u}\|+\|\mathbf{u}^h-\mathbf{v}^h\|.$$

Taylor expansion yields $\|\mathbf{u}^h-\mathbf{u}\|=O(h^2)$. Then the proof is completed by Lemma 9.51. ☐

9.5 Problems

9.5.1 Programming Assignments

Write a C++ package to implement the one-dimensional multigrid method discussed in this chapter to solve the model problem in Definition 9.6.

I. Your package must give the user the following options:

(a) boundary conditions: Dirichlet, Neumann, or mixed (partly Dirichlet and partly Neumann).
(b) restriction operators: full weighting and injection;
(c) interpolation operators: linear and quadratic;
(d) cycles: V-cycle and FMG;

(e) stopping criteria: the number of maximum iterations and the relative accuracy ϵ of the solution;

(f) the initial guess.

As for the bottom solver, you can either implement a Gaussian elimination in your own package or use the one in BLAS or LAPACK.

II. For the function in (7.91) derive the corresponding $f(x)$ and the boundary conditions. For $\epsilon = 10^{-8}$ and the zero-vector initial guess, test your multigrid solver for all combinations of (b,c,d) in II on grids with $n = 32, 64, 128, 256$, report the residual and the reduction rate of the residuals for each V-cycle. Report the maximum norm of the error vector and the corresponding convergence rates on the four grids. You should also design at least two of your own test functions and carry out the same process.

III. Gradually reduce ϵ towards 2.2×10^{-16}, under which critical value of ϵ does your program fail to achieve the preset accuracy? Why?

The requirements III-VI in Section 7.7.1 should also be met in this assignment.