Numerical Analysis - Part II

Anders C. Hansen

Lecture 18

Iterative methods for linear algebraic systems

Linear systems in elliptic PDEs

As we have seen in the previous sections linear systems $A\mathbf{x} = \mathbf{b}$, where A is a real symmetric positive (negative) definite matrix, frequently occur in numerical methods for solving elliptic partial differential equations.

Poisson's equation on a square

A typical example we already encountered is Poisson's equation on a square where the *five-point formula* approximation yields an $n \times n$ system of linear equations with $n = m^2$ unknowns $u_{p,q}$:

$$u_{p-1,q} + u_{p+1,q} + u_{p,q-1} + u_{p,q+1} - 4u_{p,q} = h^2 f(ph, qh) \qquad (1)$$

In the *natural ordering*, when the grid points are arranged by columns, A is the following block tridiagonal matrix:

$$A = \begin{bmatrix} B & I & & & \\ I & B & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & B & I \\ & & & I & B \end{bmatrix}, \qquad B = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 1 & -4 \end{bmatrix}. \qquad (2)$$

The matrix A is symmetric and negative definite

Lemma 1

For any ordering of the grid points, the matrix A of the system (1) is symmetric and negative definite.

Proof II

Proof. Let U be any linear operator changing the grid ordering. Then U is clearly unitary ($\|Ux\|_2 = \|x\|_2$ for any x). Note that any matrix $\tilde{A}$ representing the the system of equations (1) can be written as $\tilde{A} = UAU^*$ for some unitary matrix U, where A is as in (2). Self-adjointness is preserved by unitary operators, and so is the spectrum. Thus, $\tilde{A}$ is self-adjoint (symmetric as it is real). Moreover, $\sigma(A)$ does not intersect the positive half plane by the Gershgorin theorem, so we only need to show that $0 \notin \sigma(A)$. If $Ax = 0$ then, by the definition of A, x must have elements of equal modulus, however, then the definition of B (that gives A) implies that $x = 0$. □

Poisson's equation on a square

Note that when p or q is equal to 1 or m, then the values $u_{0,q}$, $u_{p,0}$ or $u_{p,m+1}$, $u_{m+1,q}$ are known boundary values and they should be moved to the right-hand side, thus leaving fewer unknowns on the left.

For any ordering of the grid points (ph, qh) we have shown in Lemma 1 that the matrix A of this linear system is symmetric and negative definite.

Corollary 2

For the linear system (1), for any ordering of the grid, both Jacobi and Gauss-Seidel methods converge.

Proof. By Lemma 1, A is symmetric and negative definite, hence convergence of Gauss-Seidel. To prove convergence of the Jacobi method, we need negative definiteness of the matrix $2D - A$, and that follows by the same arguments as in Lemma 1: recall that the proof operates with the modulus of the off-diagonal elements and does not depend on their sign. □

Relaxation

It is often possible to improve the efficiency of the splitting method by *relaxation*. Specifically, instead of letting $(A - B)\mathbf{x}^{(k+1)} = -B\mathbf{x}^{(k)} + \mathbf{b}$, we let

$$(A-B)\widehat{\mathbf{x}}^{(k+1)} = -B\mathbf{x}^{(k)}+\mathbf{b}, \quad \text{and then} \quad \mathbf{x}^{(k+1)} = \omega\widehat{\mathbf{x}}^{(k+1)}+(1-\omega)\mathbf{x}^{(k)}$$

where ω is a real constant called the *relaxation parameter*.

Note that $\omega = 1$ corresponds to the standard "unrelaxed" iteration. Good choice of ω leads to a smaller spectral radius of the iteration matrix (compared with the "unrelaxed" method), and the smaller the spectral radius, the faster the iteration converges.

Relaxation

To this end, let us express the relaxation iteration matrix H_ω in terms of $H = -(A-B)^{-1}B$. We have

$$\widehat{\mathbf{x}}^{(k+1)} = H\mathbf{x}^{(k)} + \mathbf{v} \quad \Rightarrow \quad \begin{aligned} \mathbf{x}^{(k+1)} &= \omega\widehat{\mathbf{x}}^{(k+1)} + (1-\omega)\mathbf{x}^{(k)} \\ &= \omega H\mathbf{x}^{(k)} + (1-\omega)\mathbf{x}^{(k)} + \omega\mathbf{v}, \end{aligned}$$

hence

$$H_\omega = \omega H + (1-\omega)I.$$

It follows that the spectra of H_ω and H are related by the rule $\lambda_\omega = \omega\lambda + (1-\omega)$, therefore one may try to choose $\omega \in \mathbb{R}$ to minimize

$$\rho(H_\omega) = \max\{|\omega\lambda + (1-\omega)| : \lambda \in \sigma(H)\}.$$

Relaxation

In general, $\sigma(H)$ is unknown, but often we have some information about it which can be utilized to find a "good" (rather than "best") value of ω. For example, suppose that it is known that $\sigma(H)$ is real and resides in the interval $[\alpha, \beta]$ where $-1 < \alpha < \beta < 1$. In that case we seek ω to minimize

$$\max\{|\omega\lambda + (1-\omega)| : \lambda \in [\alpha, \beta]\}.$$

Relaxation

It is readily seen that, for a fixed $\lambda < 1$, the function $f(\omega) = \omega\lambda + (1-\omega)$ is decreasing, therefore, as ω increases (decreases) from 1 the spectrum of H_ω moves to the left (to the right) of the spectrum of H. It is clear that the optimal location of the spectrum $\sigma(H_\omega)$ (or of the interval $[\alpha_\omega, \beta_\omega]$ that contains $\sigma(H_\omega)$) is the one which is centralized around the origin:

$$-[\omega\alpha+(1-\omega)] = \omega\beta+(1-\omega) \quad \Rightarrow \quad \begin{aligned} \omega_{\mathrm{opt}} &= \frac{2}{2-(\alpha+\beta)} \\ -\alpha_{\omega_{\mathrm{opt}}} &= \beta_{\omega_{\mathrm{opt}}} = \frac{\beta-\alpha}{2-(\alpha+\beta)}\,. \end{aligned}$$

Attenuation for different frequencies

The speed of convergence of some iterative methods (Jacobi with relaxation, Gauss–Seidel, etc.) can be increased drastically within the context of solving linear equations that originate in the discretization of PDEs. Herewith we analyse (with a great deal of hand-waving) the system $A\mathbf{u} = \mathbf{b}$ originated from the 5-point formula for the Poisson equation on an $m \times m$ square grid Ω_h, being solved by the *damped* Jacobi iteration.

The damped Jacobi iteration

This is the Jacobi method with a relaxation parameter ω:

$$\begin{aligned} \widehat{\mathbf{u}}^{(\nu+1)} &= -D^{-1}(A-D)\mathbf{u}^{(\nu)} + D^{-1}\mathbf{b} &= (I - D^{-1}A)\mathbf{u}^{(\nu)} + D^{-1}\mathbf{b} \\ \mathbf{u}^{(\nu+1)} &= \omega\widehat{\mathbf{u}}^{(\nu+1)} + (1-\omega)\mathbf{u}^{(\nu)} &= (I - \omega D^{-1}A)\mathbf{u}^{(\nu)} + \omega D^{-1}\mathbf{b}\,. \end{aligned}$$

The error decay is expressed in terms of the iteration matrix H_ω:

$$\mathbf{e}^{(\nu)} = [H_\omega]^\nu \mathbf{e}^{(0)}, \qquad H_\omega = I - \omega D^{-1}A = I + \tfrac{1}{4}\omega A\,,$$

and it follows from the results of Lecture 2 that the eigenvectors and the eigenvalues of H_ω are

$$\mathbf{w}^{k,\ell} = (\sin ix \sin jy), \qquad \lambda_{k,\ell}(\omega) = 1 - \omega\Big(\sin^2 \tfrac{x}{2} + \sin^2 \tfrac{y}{2}\Big),$$

$$x = \tfrac{k\pi}{m+1}, \quad y = \tfrac{\ell\pi}{m+1}.$$

The damped Jacobi iteration

We see that $\rho(H_\omega) < 1$ for any ω in $(0, 1]$, guaranteeing convergence, although a very slow one. In particular, for the "pure" Jacobi iteration (with $\omega = 1$) we have $\rho(H_J) = 1 - 2\sin^2 \frac{\pi}{2(m+1)} \approx 1 - \frac{\pi^2}{2m^2}$, and for $\omega < 1$ the spectral radius is even closer to 1.

The damped Jacobi iteration

However, expanding the error with respect to the (orthogonal) eigenvectors we obtain

$$\mathbf{e}^{(\nu)} = \sum_{k,\ell} a_{k,\ell}^{(\nu)} \mathbf{w}^{k,\ell}, \quad \mathbf{e}^{(\nu)} = [H_\omega]^\nu \mathbf{e}^{(0)} \quad \Rightarrow \quad |a_{k,\ell}^{(\nu)}| = |\lambda_{k,\ell}(\omega)|^\nu |a_{k,\ell}^{(0)}|\,,$$

i.e. the components of $\mathbf{e}^{(\nu)}$ (with respect to the basis of eigenvectors) decay at a different rate for different frequences (k, ℓ). To this end, we define

Ω_h-low frequences (LF):

$$\mathbf{w}^{(k,\ell)} = (\sin \pi k\phi \sin \pi \ell \psi)\big|_{\Omega_h} \text{ with both } k \text{ and } \ell \text{ from } [1, \frac{m+1}{2}),$$

Ω_h-high frequences (HF):

$$\mathbf{w}^{(k,\ell)} = (\sin \pi k\phi \sin \pi \ell \psi)\big|_{\Omega_h} \text{ with either } k \text{ or } \ell \text{ from } [\frac{m+1}{2}, m].$$

The damped Jacobi iteration

Let us determine the least factor $\mu(\omega)$ by which the amplitudes of HF components are damped per iteration. We have

$$\begin{aligned}
\mu(\omega) &= \max\{\,|\lambda_{k,\ell}(\omega)| : \quad \tfrac{m+1}{2} \le k \le m, \quad 1 \le \ell \le m\} \\
&= \max\{\,|1 - \omega\,(\sin^2\tfrac{x}{2} + \sin^2\tfrac{y}{2})| : \quad \tfrac{\pi}{2} \le x \le \pi, \quad 0 \le y \le \pi\} \\
&= \max\{\,|1 - \tfrac{1}{2}\,\omega|, |1 - 2\omega|\}\,,
\end{aligned}$$

and it is seen that the optimal factor μ_* is attained when $1 - \frac{1}{2}\,\omega = -(1 - 2\omega)$, i.e. for $\omega_* = \frac{4}{5}$, and its value is $\mu_* = \frac{3}{5}$.

Low to high frequency

Therefore, for the coefficients at the HF components of $\mathbf{e}^{(\nu)}$ we obtain

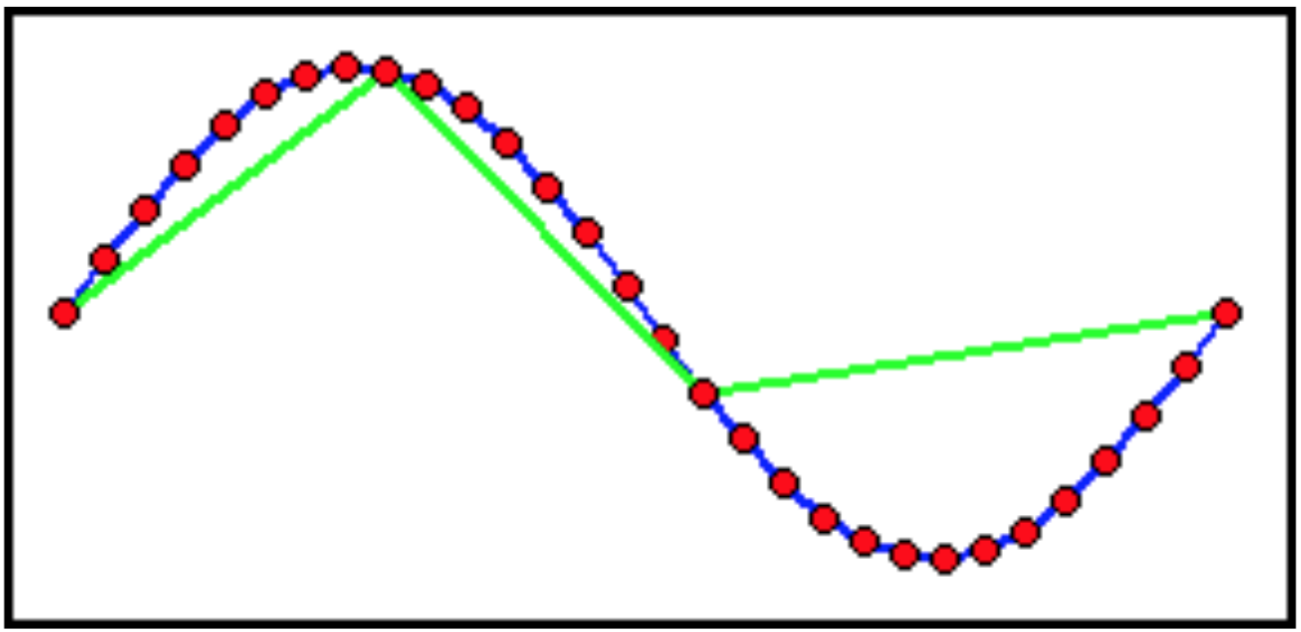
$$|a_{k,\ell}^{(\nu)}| \leq |\mu_*|^\nu |a_{k,\ell}^{(0)}| = \left(\frac{3}{5}\right)^\nu |a_{k,\ell}^{(0)}| \ll |a_{k,\ell}^{(0)}|\,,$$

i.e. the damped Jacobi method converges fast for high frequencies. For the remaining Ω_h-low frequences we notice that

$$k,\ell \in [1, \tfrac{m+1}{2}) = [1, \tfrac{1}{2h}) \quad \Rightarrow$$

$$\underbrace{\{\mathbf{w}_h^{(k,\ell)} = (\sin \pi k\phi \sin \pi \ell\psi)\big|_{\Omega_h}\}}_{\Omega_h\text{-low frequences}} \approx \underbrace{\{\mathbf{w}_{2h}^{(k,\ell)} = (\sin \pi k\phi \sin \pi \ell\psi)\big|_{\Omega_{2h}}\}}_{\Omega_{2h}\text{-high frequences}}$$

Low frequency to high frequency



Low Frequency Error (blue) on the Fine Grid Converted to High Frequency Error (green) on the Coarse Grid.

The multigrid method

The idea of the multigrid method is that, although the global error may decrease slowly by iteration, its components with high frequencies relative to Ω_h are suppressed (or smoothed) very quickly, and that dealing with the remaining components (with low frequences relative to Ω_h) we can move to the coarse grid Ω_{2h}, where these components (in part) would be of high frequencies, and thus they can be smoothed in a similar way.

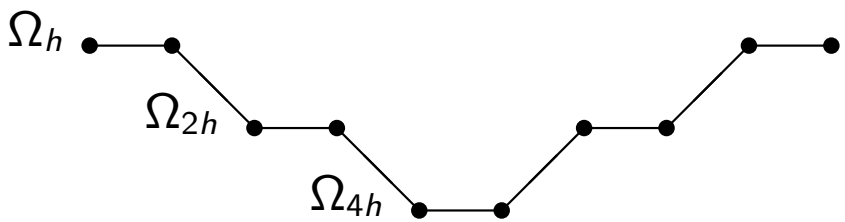
The multigrid method

Therefore, we cover the square domain by a range of nested grids, of increasing coarseness, say,

$$\Omega_h \subset \Omega_{2h} \subset \Omega_{4h} \subset \cdots \subset \Omega_{2^m h}\,.$$

At every Ω_{h_i}, the iterations (damped Jacobi, or Gauss-Seidel) remove the high frequencies relative to this grid, and we move to Ω_{2h_i}. On the coarsest grid, where the number of variables is small, we can afford to solve the equations with a direct method, by Cholesky, say.

A *multigrid sweep* starts at the finest grid, travels to the coarsest (where we apply a direct solver) and back to the finest:



Example of finer and coarser grids

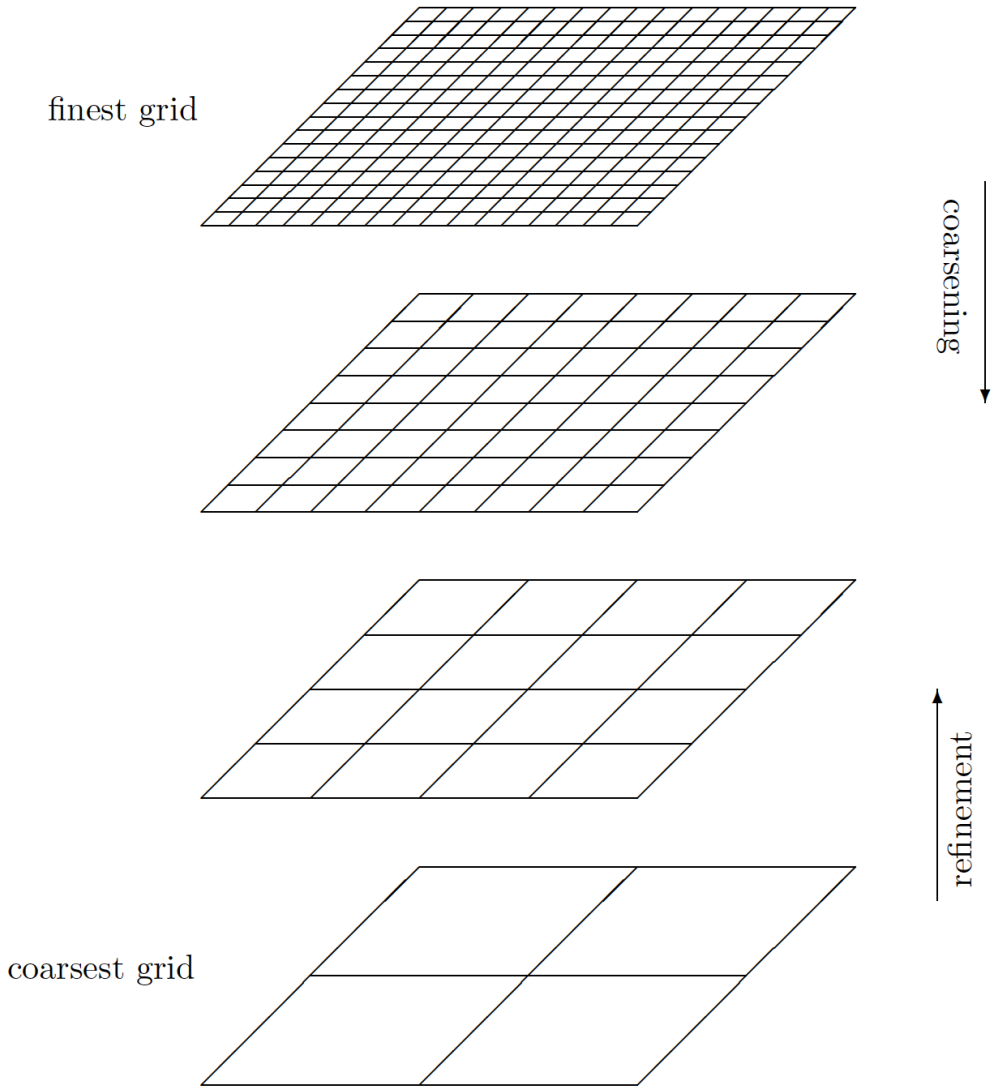
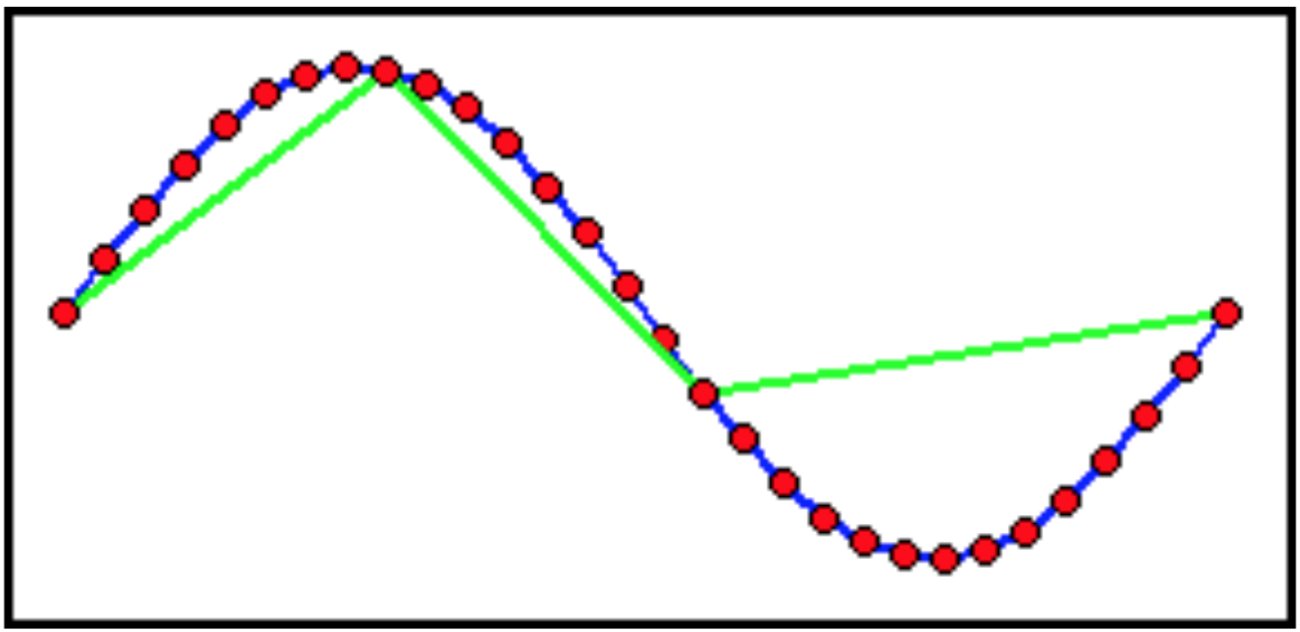


Figure : Nested grids, from the finest to the coarsest

Low frequency to high frequency



Low Frequency Error (blue) on the Fine Grid Converted to High Frequency Error (green) on the Coarse Grid.

Multigrid in practice

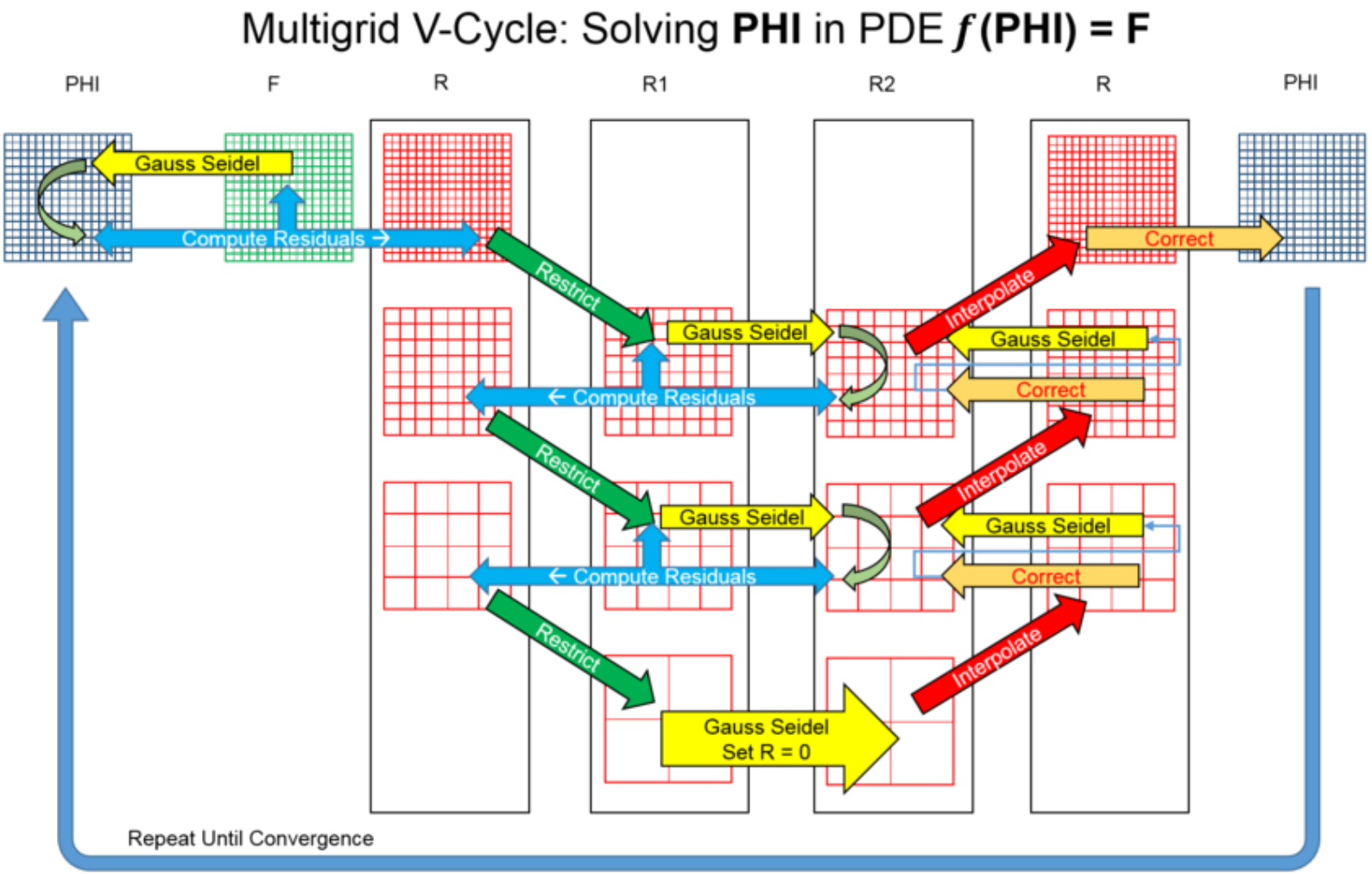


Figure : Wikipedia's interpretation of multigrid

The multigrid method

1) Each *coarsening* stage involves few (ν) iterations, then computing the residual $\mathbf{r}_h^{(\nu)} = \mathbf{b}_h - A_h\mathbf{u}_h^{(\nu)}$ (where h is the size of the grid Ω_h) and restricting it to the coarser grid Ω_{2h} via "restriction" mapping $R : \mathbb{R}^N \to \mathbb{R}^{N/4}$ as $\mathbf{r}_{2h}^{(0)} = R\mathbf{r}_h^{(\nu)}$:

$$\mathbf{r}_h^{(\nu)} = \mathbf{b}_h - A_h\mathbf{u}_h^{(\nu)} = A_h\mathbf{e}_h^{(\nu)} = \sum_{1\le k<m+1}\sum_{1\le \ell<m+1} a_{k,\ell}^{(\nu)}\mathbf{w}_h^{k,\ell}$$

$$= \sum_{1\le k<\frac{1}{h}}\sum_{1\le \ell<\frac{1}{h}} a_{k,\ell}^{(\nu)}\sin\pi k\phi\sin\pi\ell\psi\Big|_{\Omega_h}$$

$$= \underbrace{\sum_{1\le k<\frac{1}{2h}}\sum_{1\le \ell<\frac{1}{2h}} a_{k,\ell}^{(\nu)}\sin\pi k\phi\sin\pi\ell\psi\Big|_{\Omega_h}}_{\downarrow} + \underbrace{\left(\sum_{\frac{1}{2h}\le k<\frac{1}{h}}\sum_{1\le \ell<\frac{1}{h}} + \sum_{1\le k<\frac{1}{h}}\sum_{\frac{1}{2h}\le \ell<\frac{1}{h}}\right) a_{k,\ell}^{(\nu)}\sin\pi k\phi\sin\pi\ell\psi\Big|_{\Omega_h}}_{\downarrow}$$

transferred (restricted) to Ω_{2h} as

$$\sum_{1\le k<\frac{1}{2h}}\sum_{1\le \ell<\frac{1}{2h}} a_{k,\ell}^{(\nu)}\sin\pi k\phi\sin\pi\ell\psi\Big|_{\Omega_{2h}} +$$

suppressed to $\boldsymbol{\epsilon}_h^{(\nu)}$ and transferred to Ω_{2h} as

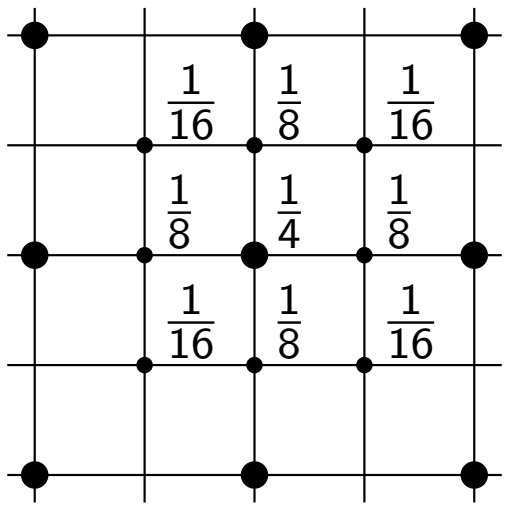
$$\boldsymbol{\epsilon}_{2h}^{(\nu)} =: \mathbf{r}_{2h}^{(0)}$$

Other helpful notes

https://ocw.mit.edu/courses/mathematics/18-086-mathematical-methods-for-engineers-ii-spring-2006/readings/am63.pdf

The multigrid method - General restriction map

A typical restriction operator R combines nine "fine" values around the "coarse" one according to the rule:



The multigrid method - Solving on the coarser grid

At the coarse grid Ω_{2h}, we are solving *for the residual*, i.e. we iterate μ times for (or apply a direct solver to) the low-dimensional system

$$A_{2h}\mathbf{v}_{2h} = 4\,\mathbf{r}_{2h}^{(0)}.$$

The factor 4 on the right is due to the fact that our linear system is of the form $A_h\mathbf{u}_h = h^2\mathbf{f}_h$, hence its coarse version is $A_{2h}\mathbf{v}_{2h} = (2h)^2\mathbf{g}_{2h}$. Having found approximative (or exact solution) $\mathbf{v}_{2h}^{(\mu)}$ we move back to Ω_h.

2) Refinement entails a prolongation via mapping $P : \mathbb{R}^{N/4} \to \mathbb{R}^N$ as

$$\mathbf{v}_h^{(\mu)} = P\mathbf{v}_{2h}^{(\mu)},$$

e.g. by linear interpolation (the exact opposite of the above procedure), and correction

$$\mathbf{u}_h^{\text{new}} = \mathbf{u}_h^{(\nu)} + \mathbf{v}_h^{(\mu)}\,.$$

The resulting vector is close to the exact solution,

$$\mathbf{u}_h^{(\nu)} + PA_{2h}^{-1}R\mathbf{r}_h^{(\nu)} \approx \mathbf{u}_h^{(\nu)} + A_h^{-1}(b_h - A_h\mathbf{u}_h^{(\nu)}) = \mathbf{u}_h^*\,.$$

The multigrid method - Sweeping through the grids

It is usual to employ only a moderate number of iterations in each restriction (3–5, say) and prolongation (just 1–2 iterations, to take care of high frequencies that have been reintroduced by prolongation) and to check for convergence only by the end of the sweep. Unless convergence occurs, we embark on another multigrid sweep and so on.