# Multigrid Methods

## Convergence Proof and Analysis

# Fixed Point Iterations

First we make some general observations about fixed point iterations. Consider the equation $Ax = b$ and the splitting $A = (A - P) + P$. We get $Pu^m = (P - A)u^{m-1} + b \Leftrightarrow u^m = \left(I - P^{-1}A\right)u^{m-1} + P^{-1}b$.

Now assume we pick some matrix $C$ that may be singular. We define $u^m = (I - CA)u^{m-1} + Cb = (I - CA)u^{m-1} + CAA^{-1}b$

Let $G \equiv (I - CA) \rightarrow (I - G) = CA$. This gives the iteration

$$\begin{aligned} u^m &= Gu^{m-1} + (I - G)A^{-1}b \\ &= G^m u^0 + \left(I + G + G^2 + \cdots + G^{m-1}\right)(I - G)A^{-1}b \\ &= G^m u^0 + \left(I - G^m\right)A^{-1}b \end{aligned}$$

This requires no assumptions on $C$ or $G$, and the above holds independent of the convergence of the iteration.

# Fixed Point Iterations

(1) $u^m = Gu^{m-1} + (I - G)A^{-1}b = G^m u^0 + (I - G^m)A^{-1}b$

If $\rho(G) < 1$, then $(I - G)^{-1} \equiv (CA)^{-1}$ exists, $C^{-1}$ must exist, and the iteration converges: $(I - G)u = (I - G)A^{-1}b \Rightarrow u = A^{-1}b$.

We can define such an iteration for any $G$, and this always corresponds to some $C$ (which can be computed from $G$ and $A$).

So, if show some complicated iteration (read 'MG') is of type (1), and we show $\rho(G) < 1$, this iteration converges to the solution.

Also, $m$ iterations are equivalent to one 'm-step' iteration: $\tilde{G} = G^m$, $u^{(0)} = u^0$, $u^{(1)} = u^m$, and setting $u^{(1)} = \tilde{G}u^{(0)} + (I - \tilde{G})A^{-1}b$

# Fixed Point Iterations

Next, we consider combining multiple iterations into a single fixed point iteration. Multigrid algorithms are of this type.

We carry out a sequence of two fixed point iteration steps (where each may be multiple steps of a particular iteration)

$$u^{(1)} = G_1 u^{(0)} + \left(I - G_1\right) A^{-1} b$$
$$u^{(2)} = G_2 u^{(1)} + \left(I - G_2\right) A^{-1} b$$

Substitution now gives

$$\begin{aligned} u^{(2)} &= G_2 G_1 u^{(0)} + G_2 \left(I - G_1\right) A^{-1} b + \left(I - G_2\right) A^{-1} b \\ &= G_2 G_1 u^{(0)} + \left(I - G_2 G_1\right) A^{-1} b \end{aligned}$$

Again, the resulting iteration can be considered a (single) fixed point iteration with a new iteration matrix ($G_2 G_1$).

# The Two-Grid Algorithm

First, consider one iteration of the two grid algorithm for $L_h u_h = f$. Vector superscripts indicate the partial steps in this single iteration.

Apply $m_1$ iterations with $G_1$ to $u_h^{(0)}$: $u_h^{(1)} = G_1^{m_1} u_h^{(0)} + (I - G_1^{m_1}) L_h^{-1} f$.

(Coarse grid correction)
Compute the residual: $r_h^{(1)} = b - L_h u_h^{(1)}$.
Restrict residual to the coarse grid: $r_H^{(1)} = I_h^H r_h^{(1)}$.
Do a direct solve for the coarse grid error: Solve $L_H e_H^{(1)} = r_H^{(1)}$.
Prolongate coarse grid error to fine grid: $\hat{e}_h^{(1)} = I_H^h r_H^{(1)}$.
Update approximate solution: $u_h^{(2)} = u_h^{(1)} + \hat{e}_h^{(1)} = u_h^{(1)} + I_H^h L_H^{-1} I_h^H r_h^{(1)}$.

Apply $m_3$ iterations with $G_3$ to $u_h^{(2)}$: $u_h^{(3)} = G_3^{m_3} u_h^{(2)} + (I - G_3^{m_3}) L_h^{-1} f$.

# The Two-Grid Algorithm

The coarse grid correction can be written in a single step as

$$u_h^{(2)} = u_h^{(1)} + C_h \left( f - L_h u_h^{(1)} \right) = \left( I - C_h L_h \right) u_h^{(1)} + C_h f,$$

where $C_h = I_H^h L_H^{-1} I_h^H$.

This can also be written as

$$u_h^{(2)} = \left( I - C_h L_h \right) u_h^{(1)} + C_h L_h L_h^{-1} f \rightarrow u_h^{(2)} = G_2 u_h^{(1)} + \left( I - G_2 \right) L_h^{-1} f,$$

where $G_2 = \left( I - C_h L_h \right) = \left( I - I_H^h L_H^{-1} I_h^H L_h \right)$.

This is again in the form of a standard fixed point iteration.

Note that since $C_h$ is singular this iteration will in general not be convergent.

# The Two-Grid Algorithm

The combination of the fixed point iterations gives the following.

$$u_h^{(1)} = G_1^{m_1} u_h^{(0)} + \left(I - G_1^{m_1}\right) L_h^{-1} b$$

Substitution into the second iteration gives

$$\begin{aligned} u_h^{(2)} &= G_2 u_h^{(1)} + \left(I - G_2\right) L_h^{-1} b \\ &= G_2 G_1^{m_1} u_h^{(0)} + G_2 \left(I - G_1^{m_1}\right) L_h^{-1} b + \left(I - G_2\right) L_h^{-1} b \\ &= G_2 G_1^{m_1} u_h^{(0)} + \left(I - G_2 G_1^{m_1}\right) L_h^{-1} b \end{aligned}$$

Substitution into the third iteration then gives

$$\begin{aligned} u_h^{(3)} &= G_3^{m_3} u_h^{(2)} + \left(I - G_3^{m_3}\right) L_h^{-1} b \\ &= G_3^{m_3} G_2 G_1^{m_1} u_h^{(0)} + \left(I - G_3^{m_3} G_2 G_1^{m_1}\right) L_h^{-1} b \end{aligned}$$

# The Two-grid Algorithm

The iteration $u_h^{(3)} = G_h^{m_3} G_2 G_h^{m_1} u_h^{(0)} + \left(I - G_h^{m_3} G_2 G_h^{m_1}\right) L_h^{-1} f$ represents the entire two-grid algorithm.
We refer to $M_h^H \equiv G_h^{m_3} G_2 G_h^{m_1}$ as the two-grid operator.
One iteration with $M_h^H$ gives for the error: $e_h^{(3)} = M_h^H e_h^{(0)}$.

For fast convergence we need to make $\rho\left(M_h^H\right)$ very small (at the least $\rho\left(M_h^H\right) < 1$ for convergence). As we have seen previously, we are mainly concerned with the smoothing factor (damping of the high frequencies), because the coarse grid correction (with exact solve) takes care of the low frequencies.

In general, it is not hard to reduce high frequency error significantly using (variations of) standard fixed point iterations.

# From Two-grid to Multigrid

Now, we define the multigrid algorithm recursively. We refer to the coarsest level as level $0$, the next finer level as level $1$, and so on till the finest level, level $l$.

Given $\gamma$, define the multigrid operator at level $j$, $M_j$, as follows.

$M_1 \equiv M_1^0$ (two grid operator – includes pre- and postsmoothing – )

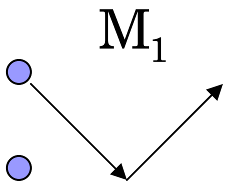
$M_2$: pre-smoothing, restrict residual to level $1$, apply $M_1$ $\gamma$ times to the residual equation, prolongate approximate error to level 2, correct approximate solution, post-smoothing.

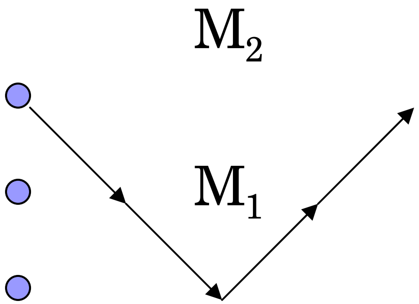
$M_3$: analogous (apply $M_2$ $\gamma$ times)

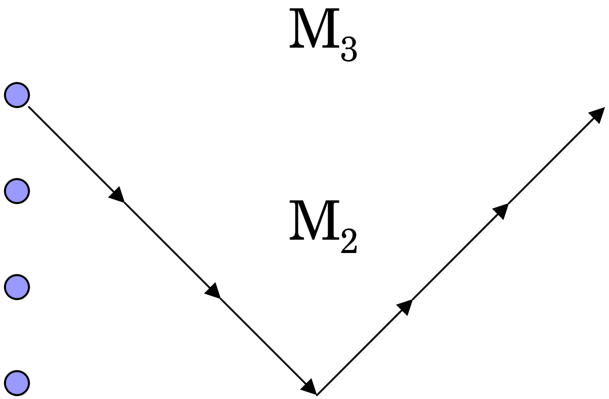
Etc.

This gives the following types of cycles for $\gamma = 1, 2$ (V or W cycle).

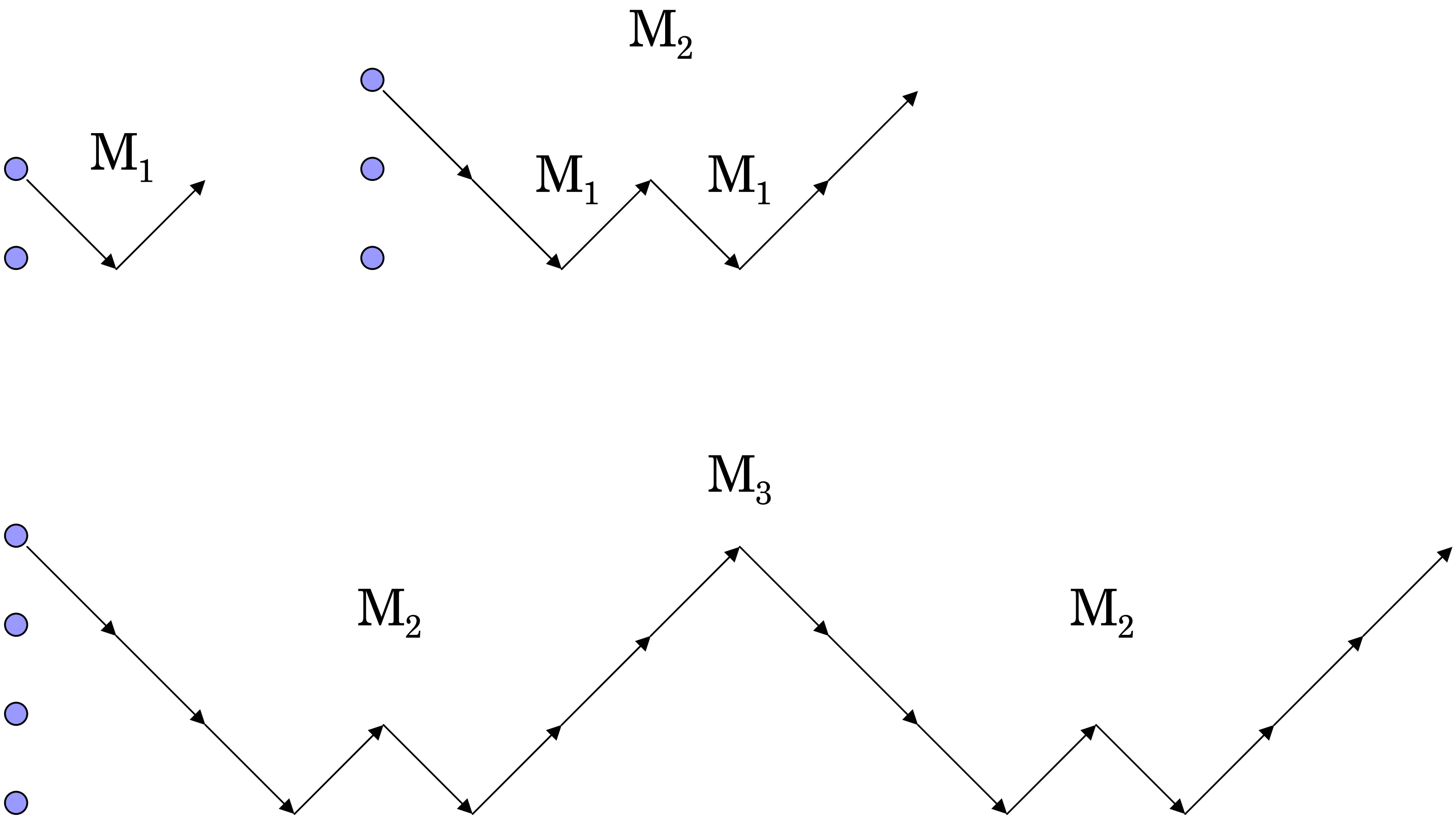
# Multigrid Cycles: $\gamma = 1$ (V-Cycle)







# Multigrid Cycles: $\gamma = 2$ (W-cycle)



# From Two-grid to Multigrid

The two-grid algorithm is the main vehicle for designing multigrid algorithms. This is important, because analyzing a multigrid algorithm (over many levels) as a whole is obviously very complicated.

It is intuitive that replacing the direct solve in the two-grid algorithm by an accurate solve using (multiple cycles of) multigrid leads to an algorithm that behaves almost like the two-grid algorithm.

We can make this precise by using the previously derived results to show that replacing the direct solve by multigrid leads to a perturbation of the two-grid operator. If the two-grid operator has small norm or spectral radius and the perturbation is small we can show that the multigrid operator converges and has almost the same rate of convergence as the two-grid algorithm.

# From Two-grid to Multigrid

Now, replace the direct solve in the two-grid algorithm at level $l$ $(u_h)$ by $\gamma$ cycles of the multigrid algorithm at level $l-1$ $(u_H)$, denoted by the operator $M_{l-1}$ (precise definition follows). The earlier discussion shows the following holds independent of the choice of $M_{l-1}$.
The coarse grid correction now runs as follows.

Compute residual and restrict: $r_l^{(1)} = f - L_l u_l^{(1)}$; $r_{l-1}^{(1)} = I_l^{l-1} r_l^{(1)}$.
Approximate solve of coarse grid equation, $L_{l-1} e_{l-1}^{(1)} = r_{l-1}^{(1)}$:

$\gamma$ iterations with $M_{l-1}$: $\hat{e}_{l-1}^{(1)} = M_{l-1}^{\gamma} \left( \hat{e}_{l-1}^{(1)} \right)_{\text{init}} + \left( I - M_{l-1}^{\gamma} \right) L_{l-1}^{-1} r_{l-1}^{(1)}$

Taking $\left( \hat{e}_{l-1}^{(1)} \right)_{\text{init}} = 0$ gives $\hat{e}_{l-1}^{(1)} = \left( I - M_{l-1}^{\gamma} \right) L_{l-1}^{-1} r_{l-1}^{(1)}$

Prolongate coarse grid correction to fine grid: $\hat{e}_l^{(1)} = I_{l-1}^{l} \hat{e}_{l-1}^{(1)}$.
Update solution: $u_l^{(2)} = u_l^{(1)} + \hat{e}_l^{(1)} = u_l^{(1)} + I_{l-1}^{l} \left( I - M_{l-1}^{\gamma} \right) L_{l-1}^{-1} I_l^{l-1} r_l^{(1)}$.

# From Two-grid to Multigrid

We rewrite $u_l^{(2)} = u_l^{(1)} + I_{l-1}^l \left(I - M_{l-1}^\gamma\right) L_{l-1}^{-1} I_l^{l-1} r_l^{(1)}$ as follows

$u_l^{(2)} = u_l^{(1)} + \tilde{C}_l \left(f - L_l u_l^{(1)}\right)$, where $\tilde{C}_l = I_{l-1}^l \left(I - M_{l-1}^\gamma\right) L_{l-1}^{-1} I_l^{l-1}$.

$$\begin{aligned} u_l^{(2)} &= \left(I - \tilde{C}_l L_l\right) u_l^{(1)} + \tilde{C}_l f = \left(I - \tilde{C}_l L_l\right) u_l^{(1)} + \tilde{C}_l L_l L_l^{-1} f \\ &= \tilde{G}_2 u_l^{(1)} + \left(I - \tilde{G}_2\right) L_l^{-1} f \end{aligned}$$

In combination with pre- and post-smoothing the multigrid iteration at level $l$ becomes

$u_l^{(3)} = G_1^{m_1} \tilde{G}_2 G_3^{m_3} u_l^{(0)} + \left(I - G_1^{m_1} \tilde{G}_2 G_3^{m_3}\right) L_l^{-1} f$,

and the multigrid operator is defined as $M_l = G_1^{m_1} \tilde{G}_2 G_3^{m_3}$.

Note that $\tilde{C}_l = C_l - I_{l-1}^l M_{l-1}^\gamma L_{l-1}^{-1} I_l^{l-1}$

# From Two-grid to Multigrid

Now we will define the multigrid operator recursively. At level $1$ (one level above the coarsest grid) we define the multigrid operator as the two-grid operator.

$$M_1 \equiv M_1^0 = G_{1,1}^{m_1} G_{2,1} G_{3,1}^{m_3} = G_{1,1}^{m_1}\left(I_1 - I_0^1 L_0^{-1} I_1^0 L_1\right) G_{3,1}^{m_3}$$

At the higher levels we define the multigrid operator in terms of the multigrid operator at the next lower level.

$$\begin{aligned} M_k &= G_{1,k}^{m_1} \tilde{G}_{2,k} G_{3,k}^{m_3} = G_{1,k}^{m_1}\left(I_k - \tilde{C}_k L_k\right) G_{3,k}^{m_3} \\ &= G_{1,k}^{m_1}\left(I_k - I_{k-1}^k \left(I - M_{k-1}^{\gamma}\right) L_{k-1}^{-1} I_k^{k-1} L_k\right) G_{3,k}^{m_3} \end{aligned}$$

If we take $M_0 = 0$ the definition of $M_k$ also holds for $k = 1$.

# The Multigrid Algorithm

Now we would like to derive convergence rates for the multigrid algorithm from the two-grid algorithm (and prove convergence).

We do this by assuming we get a good convergence rate for the two-grid algorithm (small norm or spectral radius), and then use this to derive bounds on the spectral radius or norm of the multigrid algorithm.

For level $1$ this is easy!

For level $k > 1$ we consider the multigrid operator as a perturbation of the two-grid operator at that level.

# The Multigrid Algorithm

Two-grid operator ($k$): $M_k^{k-1} = G_{1,k}^{m_1} G_{2,k} G_{3,k}^{m_3} = G_{1,k}^{m_1} \left( I_k - C_k L_k \right) G_{3,k}^{m_3}$

Multigrid operator ($k$): $M_k = G_{1,k}^{m_1} \tilde{G}_{2,k} G_{3,k}^{m_3} = G_{1,k}^{m_1} \left( I_k - \tilde{C}_k L_k \right) G_{3,k}^{m_3}$

Note that $\tilde{C}_k = C_k - I_{k-1}^k M_{k-1}^{\gamma} L_{k-1}^{-1} I_k^{k-1}$, and so

$$\begin{aligned} M_k &= G_{1,k}^{m_1} \left( I_k - \tilde{C}_k L_k \right) G_{3,k}^{m_3} \\ &= G_{1,k}^{m_1} \left( I_k - C_k L_k + I_{k-1}^k M_{k-1}^{\gamma} L_{k-1}^{-1} I_k^{k-1} L_k \right) G_{3,k}^{m_3} \\ &= G_{1,k}^{m_1} \left( I_k - C_k L_k \right) G_{3,k}^{m_3} + G_{1,k}^{m_1} I_{k-1}^k M_{k-1}^{\gamma} L_{k-1}^{-1} I_k^{k-1} L_k G_{3,k}^{m_3} \\ &= M_k^{k-1} + A_{k-1}^k M_{k-1}^{\gamma} A_k^{k-1} \end{aligned}$$

# The Multigrid Algorithm

Now assume that analysis of the two-grid algorithm (using (local) Fourier analysis, analysis of the grid transfer operators, etc.) gives

$\left\| M_k^{k-1} \right\| < \sigma$ for $k = 1, 2, \ldots, l$.

Furthermore, let $\left\| A_{k-1}^{k} \right\| \left\| A_k^{k-1} \right\| < C,$ for $k = 1, 2, \ldots, l$

This gives for the multigrid operators

$\left\| M_1 \right\| = \left\| M_1^0 \right\| < \sigma$

$$\begin{aligned} \left\| M_2 \right\| &= \left\| M_2^1 + A_1^2 M_1^\gamma A_1^2 \right\| \\ &\leq \sigma + \left\| A_1^2 \right\| \left\| A_2^1 \right\| \left\| M_1 \right\|^\gamma \\ &\leq \sigma + C\sigma^\gamma \end{aligned}$$

So, we can define bounds recursively …

# The Multigrid Algorithm

Let $\left\|M_k^{k-1}\right\| < \sigma$ and $\left\|A_{k-1}^{k}\right\|\left\|A_k^{k-1}\right\| < C$ for $k = 1, 2, \ldots, l$.

Then $\left\|M_k\right\| \leq \eta_k$, where
$\eta_1 = \sigma$, and $\eta_k = \sigma + C\eta_{k-1}^{\gamma}$, for $k = 2, 3, \ldots, l$.

If in addition $4C\sigma \leq 1$ and $\gamma = 2$, we get

$$\left\|M_l\right\| \leq \eta \equiv \frac{1-\sqrt{1-4C\sigma}}{2C} \leq 2\sigma, \text{ for } l \geq 1.$$

For $\gamma > 2$ we get (of course) better results.
In general, $C \geq 1$, but not much larger. For $\gamma = 1$ we do not get $h$-independent convergence rates. This does not mean a V-cycle cannot be faster than a W-cycle for a given problem and algorithm.

# The Multigrid Algorithm

In contrast to many of the convergence bounds we saw earlier, these bounds are generally sharp and can be made quite small.

The proof shows (once more) that the two-grid algorithm is the cornerstone of the design of a multigrid solver.

If we can put together an efficient two-grid algorithm with a small upper bound  on the norm (or spectral radius) independent of $h$, we will get good convergence. And the convergence will not deteriorate for $h \rightarrow 0$.

For most algorithms the cost of the multigrid cycle (at the finest level) is $O(N)$, where $N$ is the number of unknowns (section 2.4.3). This means a fixed amount of work per unknown.