

Numerical Methods, Multigrid



Wim A. Mulder

Faculty of Civil Engineering and Geosciences, Department of
Geoscience and Engineering, Delft University of Technology,
Delft, The Netherlands

Shell Global Solutions International BV, Amsterdam,
The Netherlands

Definition

A multigrid method is an algorithm for the iterative solution of partial differential equations using a sequence of discretizations on multiple scales.

Introduction

The numerical solution of a partial differential equation (PDE) requires its discretization and a method to solve the resulting large system of algebraic equations. For linear equations, the resulting system is often a sparse matrix, and a direct solution method suffices if the size of the problem is modest. In three space dimensions, the computational cost of a direct solver may be too high. An iterative method that improves the accuracy of an approximate solution step by step can be a good alternative. Multigrid is an example.

The multigrid method has optimal complexity: the amount of work required to solve a problem with N unknowns is $O(N)$, meaning that it scales with the problem size N . It achieves its efficiency by employing several discretization grids for the same problem.

History

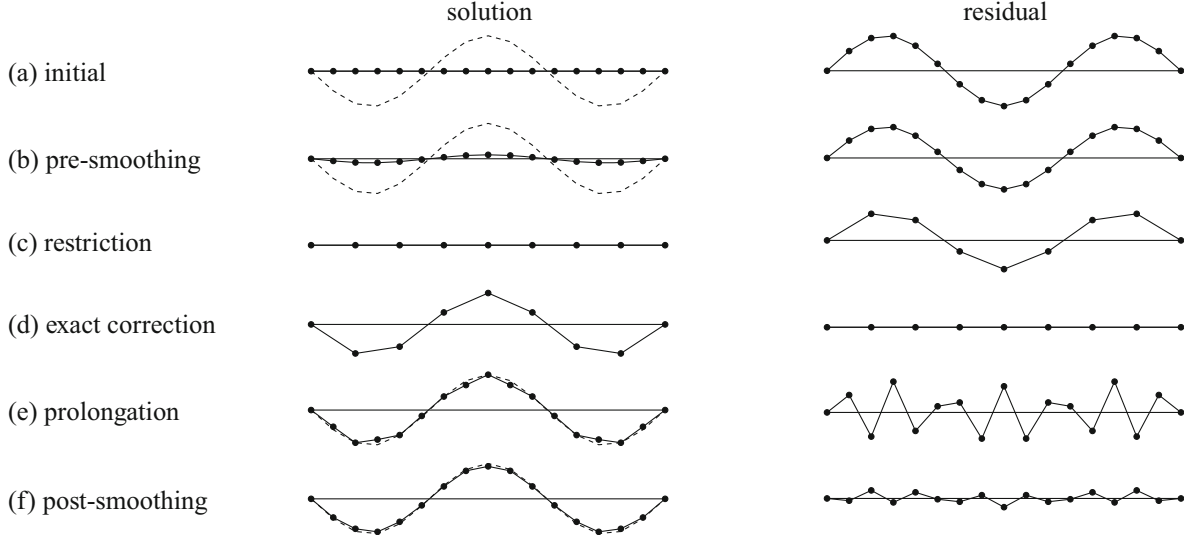
Fedorenko (1964) introduced the multigrid method as an iterative scheme for solving Poisson's equation on a square and showed that the number of computations required to determine a solution with a prescribed accuracy is proportional to the number of unknowns, N . Therefore, the method has an optimal computational complexity. Brandt (1973) found that the actual computational cost for a sufficiently accurate result was about 10 work units, where a work unit is the cost of evaluating the discretized equations. He connected the method to local adaptive grid refinement and introduced nonlinear multigrid. Hackbush (1976, 1985) discovered the method independently and provided a mathematical foundation. Since then, the method was developed further to handle PDEs other than Poisson's, which is elliptic, and to go beyond PDEs.

Textbooks include those by Hackbush (1976), who includes convergence proofs; by Wesseling (1992), with chapters on computational fluid dynamics; by Briggs et al. (2000), an easy to read introduction; and by Trottenberg et al. (2001), among others.

Two-Grid Scheme

A discrete representation of a PDE provides an algebraic system of equations with relations between solution values at neighbouring grid points. It is fairly easy to make local corrections that reduce the solution error on a short range but much harder to correct the long-range or long-wavelength components of the solution that have a more global character. By projecting the solution onto a coarser grid, the long wavelengths become shorter and can effectively be solved for. Combining the corrections to the solution from coarser and finer grids yields an efficient solver.

A simple one-dimensional example can help to understand the fundamentals of the multigrid method. The PDE is $Lu = f$,



Numerical Methods, Multigrid, Fig. 1 Steps in a two-grid cycle. (a) The initial solution on the fine grid is set to zero and the residual equals the forcing function. The dash line represents the exact numerical solution. (b) After one step of pre-smoothing with damped Jacobi relaxation, the solution error is still large, and the residual has hardly changed in this example. Its restriction to the coarser grid (c) is solved exactly, using a coarse-grid version of the discrete differential operator.

The resulting correction to the fine-grid solution (d) is interpolated or prolonged back to the fine grid and added to the fine-grid solution (e). The error in the solution is now dominated by the short wavelengths and appears as an oscillatory function, which is reflected in the corresponding residual. A post-smoothing step (f) removes most of the solution error. Repeating the whole cycle will further reduce the error

with $u(x)$ the unknown solution as a function of position x on a finite interval $[x_{\min}, x_{\max}]$ on the real axis, $f(x)$ a source term or forcing function, and L a linear differential operator, for instance, minus the Laplacian, which in 1D is $-\frac{d^2}{dx^2}$. Dirichlet boundary conditions let the solution be zero at the endpoints of the interval. To obtain a discrete representation of the problem, an equidistant one-dimensional grid is defined with grid points $x_k = x_{\min} + kh$, index $k = 0, \dots, N+1$, where $N+1 = 2^M$ and M is a positive integer. The grid spacing is $h = (x_{\max} - x_{\min})/(N+1)$. A standard second-order finite-difference scheme leads to

$$\frac{-u_{k-1} + 2u_k - u_{k+1}}{h^2} = f_k, \quad k = 1, \dots, N,$$

where u_k approximates $u(x_k)$ and $f_k = f(x_k)$. At the boundaries, $u_0 = u_{N+1} = 0$. The discrete equations represent the problem $L^h u^h = f^h$, where L^h is a $N \times N$ sparse matrix and u^h and f^h are vectors of length N . The residual is defined as $r^h = f^h - L^h u^h$ and should vanish once the numerical solution has been found.

A simple iterative method is Jacobi relaxation, in which the matrix L^h is replaced by its diagonal $D^h = 2/h^2$. One step of Jacobi relaxation amounts to

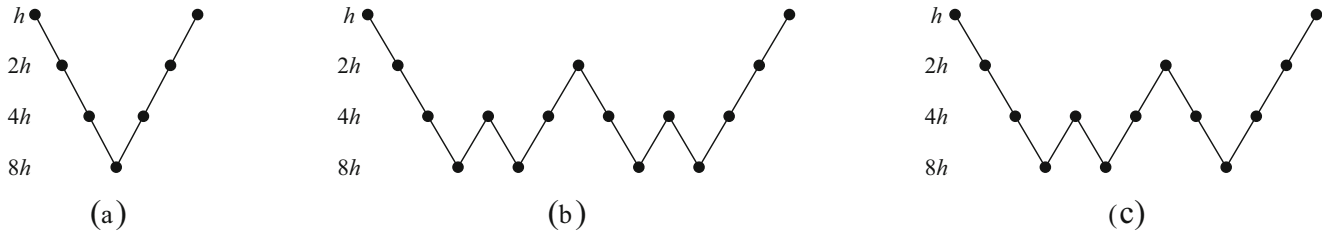
$$u^h := u^h + \omega (D^h)^{-1} r^h.$$

The symbol “:=” indicates that the solution values are replaced by the expression on the right-hand side. The factor

ω controls the amount of damping. Convergence requires $0 < \omega \leq 1$. Fourier analysis (Hackbush 1985) shows that the convergence rate, the factor by which the norm of the difference between the current and the exact numerical solution is reduced, is $1 - O(h^2)$ for Jacobi relaxation. Convergence slows down on increasingly finer grids.

The Fourier analysis reveals that the slow convergence is caused by the long-wave components of the solution. Because long waves can be represented on coarser grids, it makes sense to have look at a two-grid scheme, outlined in Fig. 1 and its caption. The coarse grid consists of every other grid point of the fine grid: $x_K = x_0 + KH$, with $H = 2h$ and $K = 0, \dots, \frac{1}{2}(N+1)$. A restriction operator $\tilde{I}_h^H$ maps the current fine-grid residual r^h to the coarser grid: $r^H = \tilde{I}_h^H r^h$. The simplest restriction operator is injection: $r_K^H = r_{2K}^h$, $K = 1, \dots, \frac{1}{2}(N-1)$. Full weighting, which lets $r_K^H = \frac{1}{4}r_{2K-1}^h + \frac{1}{2}r_{2K}^h + \frac{1}{4}r_{2K+1}^h$, applies some smoothing to the residual and is more common. The exact solution of the coarse-grid problem yields the coarse-grid correction $v^H = (L^H)^{-1} r^H$, which should be interpolated back to the fine grid and added as a correction to the latest fine-grid solution. The interpolation or prolongation operator is denoted by I_H^h and lets $u^h := u^h + I_H^h v^H$.

If linear interpolation is used for prolongation, this becomes



Numerical Methods, Multigrid, Fig. 2 (a) V-cycle, (b) W-cycle, and (c) F-cycle

$$u_{2K}^h := u_{2K}^h + v_K^H, \quad K = 1, \dots, \frac{1}{2}(N-1),$$

$$u_{2K-1}^h := u_{2K-1}^h + \frac{1}{2}(v_{K-1}^H + v_K^H), \quad K = 1, \dots, \frac{1}{2}(N+1).$$

Here, it is assumed that $v_0^H = v_{(N+1)/2}^H = 0$. After prolongation, an additional relaxation step with damped Jacobi further removes oscillatory error components of the solution. Jacobi relaxation inside this two-grid scheme has a different purpose than when used as an iterative solver by itself. Instead of removing both the short- and long-wave components of the numerical solution error, it only has to deal with those components that cannot be represented on the coarser grid without aliasing. Therefore, a damped version with $\omega < 1$ can be more effective as it can be optimized to remove the short-wave or oscillatory components. For Fig. 1, $\omega = 2/3$ was used. The optimal choice requires a more detailed analysis (Hackbush 1985).

A relaxation scheme geared towards removing oscillatory components is called a smoother. Another popular choice is Gauss-Seidel relaxation. The operator is then approximated by a lower or upper triangular matrix, which is easy to invert. The implementation of the scheme is similar to Jacobi, but the residual is evaluated with the most recent solution available. The result will depend on the order in which the grid is traversed. With lexicographic Gauss-Seidel, one follows the natural index k in increasing order or in the opposite direction. Symmetric Gauss-Seidel performs both these smoothing steps in sequence. An alternative is red-black Gauss-Seidel, where first the points with an odd and then those with an even index are updated, always using the latest solution values for the residual. In 2D, this would follow a checkerboard pattern.

The grid transfer operator, restriction and prolongation, should obey $m_p + m_r > 2m$ (Hackbush 1985) where $2m$ is the order of the differential equation and $m_p - 1$ is the highest degree of the polynomial that is interpolated exactly. The scaled adjoint of the restriction operator is an interpolation operator for which m_r can be defined in the same way as m_p . In the example above, $2m = 2$, full-weighting has $m_r = 2$, and prolongation based on linear interpolation has $m_p = 2$. More advanced grid transfer operators are based on the differential

operator L^h , leading to operator-weighted restriction and prolongation.

The coarse-grid operator L^H can be based on the same discretization as the fine-grid operator L^h . An alternative is the Galerkin coarse-grid approximation $\tilde{L}_h^H L^h I_H^h$. In the current example with full weighting, these happen to be the same. Operator-weighted grid transfer operators can accelerate the convergence by using the coefficients of the differential operator in the construction of restriction and prolongation operators, which in turn will affect the Galerkin coarse-grid approximation of the differential operator.

Multigrid

Instead of using the exact numerical solution on the coarser of the two grids, a two-grid scheme can be applied to the coarse-grid problem. Extending this recursively to a coarsest grid with three points and one unknown, a multigrid solver is obtained. All these grids can be visited in different orders. The standard approach is to start at the finest grid, perform a number of pre-smoothing steps, then move to the coarser grid, again perform pre-smoothing, and continue this until the coarsest grid is reached. There, the exact numerical solution is computed, and the resulting coarse-grid correction is prolonged back to the finer level, followed by a number of post-smoothing steps. This is repeated up to the finest grid. Such a sequence is called a V-cycle. A single V-cycle may be insufficient to obtain a convergence rate on the coarser grid that is similar to the smoothing rate on the finer grid. Convergence on coarser grids can be improved by performing more than one cycle. In the W-cycle, the number of cycles doubles on each coarser grid. Figure 2 illustrates the order in which grids are visited for a V-, W-, and F-cycle. The last one is less expensive than a W-cycle, as the number of cycles increases by only one for increasingly coarser grids. Another option, adaptive cycling, decides on the sequence of grids by monitoring the decrease of the residuals.

With a proper choice of relaxation scheme or smoother and grid transfer operators, the multigrid method can reach a grid-independent or h -independent convergence rate. This

does not, however, mean that the amount of work required to reach a sufficiently accurate solution is proportional to the number of grid points N . The exact numerical solution differs from the true solution of the PDE because of the numerical error due to the discretization, typically with some power p of h . In the earlier example, the discretization is second-order accurate, so $p = 2$. Given this numerical error, convergence to machine precision is more than needed. An iteration error somewhat smaller than the discretization error should suffice. Therefore, more iterations should be carried out if a more accurate solution on a finer grid has to be computed. This leads to an overall complexity of $O(N \log N)$ rather than $O(N)$. Successive grid refinement, in which one first computes a solution on a coarse grid and uses that as an initial guess for the next finer grid and so on, enables removal of the factor $\log N$, assuming that a fixed number of multigrid cycles are used at each level. The combination of successive grid refinement with a multigrid solver is called full multigrid (FMG).

Nonlinear Multigrid

One way to apply the multigrid method to nonlinear PDEs is the use of Newton's method. A multigrid solver is then applied to a linearization of the discretized nonlinear problem. Another approach is the use of the discretized nonlinear PDE inside the multigrid algorithm. This requires the full solution to be available on the coarser grids. In the example above, only corrections to the solution were represented on coarser grids. The full approximation scheme (FAS) is a reformulation of the multigrid method that includes the full solution (Brandt 1982). It requires a restriction of the full solution, $u^H = I_h^H u^h$, and of the coarse-grid forcing function, $f^H = L^H(I_h^H u^h) + \tilde{I}_h^H r^h$. The prolongation needs to be changed to $u^h := u^H + I_h^h(u^H - I_h^H u^h)$, where u^H is the sufficiently converged solution on the coarser grid and $I_h^H u^h$ the restriction of the latest fine-grid solution. The restriction operators for the residuals, $\tilde{I}_h^H$, and for the solution, I_h^H , do not have to be the same. Note that the coarse-grid problem, $L^H u^H = f^H$, can be interpreted in a different way: $L^H u^H = f^H = \tilde{I}_h^H f^h + \tau_h^H$. The fine-to-coarse defect correction, $\tau_h^H = L^H(I_h^H u^h) - \tilde{I}_h^H(L^h u^h)$, ensures that the coarse-grid equations maintain the accuracy of the fine-grid equations.

Nonlinear equations of the form $L^h(u^h) = f^h$ fit into this scheme. For $L^H(u^H)$, the same discretization scheme as on the fine grid can be adopted. Smoothing operators can be based on a local linearization of the residuals.

Generalizations

Multigrid is highly efficient for the solution of Poisson's equation on the square, discretized by the finite-difference method on a Cartesian grid. For other PDEs, describing convection, diffusion, waves, and flow, the method may be less easy to apply. The same is true for unstructured grids, for instance, based on triangles or tetrahedra. With suitable modifications and at the expense of additional code complexity, the multigrid method can still be the optimal approach. With unstructured grids, the coarser and finer grid are generally not nested, leading to grid transfer operations that are less easy to code than on a Cartesian grid.

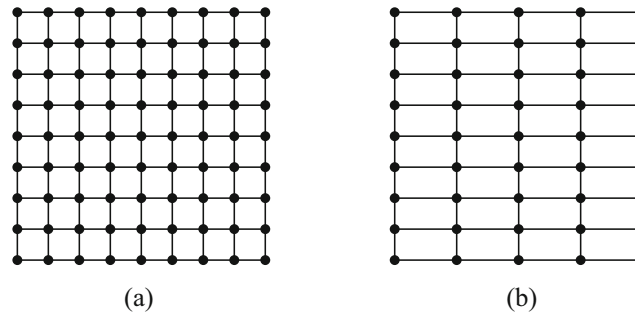
Anisotropy in the PDE can degrade performance. Consider the equation:

$$-\frac{\partial^2 u}{\partial x^2} - a \frac{\partial^2 u}{\partial y^2} = f(x, y).$$

For $0 < a \ll 1$, smoothing operators may have difficulty with the y -direction where there is only weak coupling between neighbouring unknowns. Effective anisotropy in the discretized equation can also happen with grids that have widely different spacings in the various coordinate directions. A more powerful relaxation scheme may repair the poor performance of a simple smoother. For 2D problems, line relaxation will smooth in one direction and solve in the other. An example is line Jacobi relaxation, which approximates the matrix L^h by dropping the off-diagonals in one coordinate direction, while keeping them in the other. The dropped off-diagonals can be subtracted from the main diagonal with a suitable weight factor to obtain better smoothing properties. For 3D problems, plane relaxation is required, which may be too expensive. An alternative is semi-coarsening, where the coarse grid has the same number of points in the direction of weak coupling and is coarsened only in the direction in which the smoothing effectively removes the shorter wavelengths. Figure 3 shows an example of a grid before and after semi-coarsening in the horizontal direction. In the more general case where semi-coarsening is required in all coordinates, it can be applied in alternating directions. A more costly but powerful approach is simultaneous semi-coarsening in all coordinates (Mulder 1989), which still retains optimal complexity.

Beyond Partial Differential Equations

The application of the multigrid method is not restricted to the solution of PDEs. There are generalizations to eigenvalue problems, integral equations, optimization and optimal control, statistical physics, image processing, image



Numerical Methods, Multigrid, Fig. 3 Semi-coarsening of grid (a) in the horizontal direction provides grid (b)

segmentation, and edge detection. In all cases, the simultaneous use of different scales overcomes the problem of slowly converging solution components.

Algebraic multigrid (AMG), or, better, the algebraic multi-level method, refers to a class of solvers that construct sets of coarse-level equations without referring to grids or geometrical properties (Trottenberg et al. 2001; Shapira 2008). Coarsening is applied to subsets of unknowns for which smoothing is effective. If the problem is specified by a large matrix, these subsets can be determined by examining the coefficients of the matrix. Strongly coupled unknowns, say u_i and u_j , are related by a matrix coefficient a_{ij} that is relatively large. In that case, smoothing is effective, and one of the unknowns can be removed from the equations on the coarser level. The coarser level will then be mainly populated by weakly coupled variables. Operator-weighted grid transfer operators are natural in this context, as well as the Galerkin approach for the construction of the coarse-level operator. Algebraic multilevel methods are a popular choice for finite-element discretizations on unstructured grids, as they can be used as a black-box solver and do not require the tedious coding of grid transfer operations between non-nested elements.

The multigrid method bears some resemblance to other techniques such as wavelets, hierarchical-basis finite elements, multi-scale techniques, cyclic reduction, fast multipole methods, and other divide-and-conquer methods.

Geophysical Applications

Multigrid has been applied successfully to a wide range of partial differential equations, describing static and dynamics problems, diffusion, convection, and flow problems. Examples in geophysics include elliptic problems in potential theory, such as gravity (Kusche 2002), magnetostatics (De Giersem and Hameyer 2001), electrostatics (Ascher and Haber 2003), and Darcy's law in porous media flow (Mulder and Gmelig Meyling 1993). Two-dimensional phase unwrapping for satellite interferometric synthetic

aperture radar (Pritt 1996) can also be stated as an elliptic problem. Controlled-source electromagnetics (Werthmüller et al. 2019), magnetotellurics (Li et al. 2015), mantle convection, and geodynamics (Gerya 2019) are examples of parabolic problems.

The wave equation for seismic applications is a hyperbolic problem. Iterative solution of its frequency-domain formulation, the Helmholtz equation, was a notoriously difficult numerical problem but can nowadays be accomplished by using a damped version as preconditioner in a conjugate-gradient-type iterative scheme. Multigrid efficiently deals with the approximate inversion of the preconditioner (Riyanti et al. 2006). In the generalization of the method to the elastic wave equation, simple Point-Jacobi smoothing turns out to be insufficient for the secondary or shear waves, and a more powerful smoother is required (Rizzuti and Mulder 2016).

Summary

The multigrid method provides an optimal iterative solution method for a wide class of problems governed by partial differential equations. Its application to a finite-difference discretization of Poisson's equation on a square is the easiest. For other problems, a bit more effort may be required.

Cross-References

- [Numerical Methods, Finite Difference](#)
- [Numerical Methods, Finite Element](#)

Bibliography

- Ascher UM, Haber E (2003) A multigrid method for distributed parameter estimation problems. *Electron Trans Numer Anal* 15:1–17
- Brandt A (1973) Multi-level adaptive technique (MLAT) for fast numerical solution to boundary value problems. In: Cabannes H, Temam R (eds) *Proceedings of the third international conference on*

- numerical methods in fluid mechanics. Springer Berlin Heidelberg, Berlin/Heidelberg, pp 82–89. <https://doi.org/10.1007/BFb0118663>
- Brandt A (1982) Guide to multigrid development. In: Hackbusch W, Trottenberg U (eds) Multigrid methods. Springer Berlin Heidelberg, Berlin/Heidelberg, pp 220–312. <https://doi.org/10.1007/BFb0069930>
- Briggs WL, Henson VE, McCormick SF (2000) A multigrid tutorial, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia. <https://doi.org/10.1137/1.9780898719505>
- De Gersem H, Hameyer K (2001) Full multigrid for magnetostatics using unstructured and non-nested meshes. IEEE Trans Magn 37(5):3460–3464. <https://doi.org/10.1109/20.952637>
- Fedorenko RP (1964) The speed of convergence of one iterative process. USSR Comput Math Math Phys 4(3):227–235. [https://doi.org/10.1016/0041-5553\(64\)90253-8](https://doi.org/10.1016/0041-5553(64)90253-8)
- Gerya TV (2019) Introduction to numerical geodynamic modelling, 2nd edn. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781316534243>
- Hackbusch W (1976) Ein iteratives Verfahren zur schnellen Auflösung elliptischer Randwertprobleme. Report 76-12. Mathematisches Institut, Universität zu Köln, Köln
- Hackbusch W (1985) Multi-grid methods and applications. Springer series in computational mathematics, vol 4. Springer, Berlin/Heidelberg. <https://doi.org/10.1007/978-3-662-02427-0>
- Kusche J (2002) On fast multigrid iteration techniques for the solution of normal equations in satellite gravity recovery. J Geodyn 33(1): 173–186. [https://doi.org/10.1016/S0264-3707\(01\)00062-X](https://doi.org/10.1016/S0264-3707(01)00062-X)
- Li G, Hao T, Zhang L (2015) Performance of preconditioned iterative and multigrid solvers in solving the three-dimensional magnetotelluric modeling problem using the staggered finite-difference method: a comparative study. J Geophys Eng 13(1):1–10, 12. <https://doi.org/10.1088/1742-2132/13/1/1>
- Mulder WA (1989) A new multigrid approach to convection problems. J Comput Phys 83(2):303–323. [https://doi.org/10.1016/0021-9991\(89\)90121-6](https://doi.org/10.1016/0021-9991(89)90121-6)
- Mulder WA, Gmelig Meyling RHJ (1993) Numerical simulation of two-phase flow using locally refined grids in three space dimensions. SPE Adv Technol Ser 1:36. <https://doi.org/10.2118/21230-PA>
- Pritt MD (1996) Phase unwrapping by means of multigrid techniques for interferometric SAR. IEEE Trans Geosci Remote Sens 34(3): 728–738. <https://doi.org/10.1109/36.499752>
- Riyanti CD, Erlangga YA, Plessix R-É, Mulder WA, Vuik C, Oosterlee C (2006) A new iterative solver for the time-harmonic wave equation. Geophysics 71(5):E57–E63. <https://doi.org/10.1190/1.2231109>
- Rizzuti G, Mulder WA (2016) Multigrid-based ‘shifted-Laplacian’ preconditioning for the time-harmonic elastic wave equation. J Comput Phys 317:47–65. <https://doi.org/10.1016/j.jcp.2016.04.049>
- Shapira Y (2008) Matrix-based multigrid: theory and applications. Numerical methods and algorithms, vol 2, 2nd edn. Springer, Boston. <https://doi.org/10.1007/978-0-387-49765-5>
- Trottenberg U, Oosterlee CW, Schüller A (2001) Multigrid. Academic, London. ISBN 9780127010700
- Werthmüller D, Mulder WA, Slob EC (2019) emg3d: a multigrid solver for 3D electromagnetic diffusion. J Open Source Softw 4(37):1463. <https://doi.org/10.21105/joss.01463>
- Wesseling P (1992) An introduction to multigrid methods. Wiley, Chichester