NIEUW ARCHIEF VOOR WISKUNDE (3), XXIX (1981), 71-101

INTRODUCTION TO MULTIGRID METHODS

P.W. HEMKER

A convenient way to give an introduction to multigrid methods is by means of the notion of "Defect Correction Process". Defect correction processes are general iterative processes for the approximation of operator equations. A large number of well known iterative methods can be classified into this category, and among these are the multigrid methods. Therefore, we give an introduction to elementary defect correction processes (DCP) in Section 1. In Section 2 we shall elaborate the idea of DCP to get the framework to fit the multigrid methods in. In Section 3 we give a short introduction to the discretization of analytic problems, with special emphasis on the discretization on related grids, as they are used in multigrid methods. In Section 4 we treat the principles of multigrid algorithms and we give the basic structure of convergence proofs of multigrid methods. Some examples of applications of multigrid methods are given in Section 5.

1. ELEMENTARY DEFECT CORRECTION PROCESSES

In principle, a defect correction process is an iterative process to solve an equation that we cannot or don't want to solve directly:

(P) $\qquad Fx = y,$

where F is a mapping from A to B. A and B are normed linear spaces. In general the mapping F is non-linear, F is not defined on the whole of A and F is neither injective nor surjective.

We assume that there exist subsets $X \subset A$ and $Y \subset B$ such that F is defined on the whole of X, and the mapping $F: X \to Y$ is surjective. In addition we often require that there exists a unique $x \in X$ such that

$Fx = y$ (or: in addition the mapping $F:X \to Y$ is injective and hence it is bijective). We assume that we *can* solve some approximations $(\tilde{P})$ of the problem (P), i.e. for all $\tilde{y} \in \tilde{Y} \subset Y$ we can solve the equation

$$(\tilde{P}) \qquad \tilde{F}\tilde{x} = \tilde{y}, \qquad \tilde{x} \in \tilde{X},$$

where $\tilde{F}: X \to \tilde{Y}$ is some "approximation" of the operator F.
Formally we describe this as follows:
We assume that for some subset $\tilde{Y} \subset Y$ with $y \in \tilde{Y}$, there exists a mapping

$$\tilde{G}: \tilde{Y} \to X,$$

which we shall call the *approximate inverse* of F.

The meaning of $\tilde{G}$ is that for any $y \in \tilde{Y}$ an approximation to the solution of the equation $Fx = y$ is given by $\tilde{G}y$. The mapping $\tilde{G}$ needs not to be linear and neither injective nor surjective.

REMARK 1. If $\tilde{G}$ is *not* surjective, then possibly $x \notin \tilde{G}\tilde{Y}$, with x the solution of $Fx = y$.

REMARK 2. If $\tilde{G}$ is injective, then an $\tilde{F}: \tilde{G}\tilde{Y} \to \tilde{Y}$ exists such that $\tilde{F}\tilde{G} = I_{\tilde{Y}}$, where $I_{\tilde{Y}}$ is the identity operator on $\tilde{Y}$. Then $\tilde{F}$ is the left-inverse of $\tilde{G}$ and $\tilde{F}$ is "an approximation to F". However, we notice that $\tilde{F}$ is only defined on $\tilde{G}\tilde{Y}$ and not on X!

In a Defect Correction Process the solution of the original problem (P) *is found (or approximated) by the iterative application of one (or more) approximate inverse(s)* $\tilde{G}$.

In its most elementary form we have two versions of the defect correction process for the solution of (P):
the *first defect correction process*

$$\text{DCPA} \begin{cases} x_0 = \tilde{G}y, \\ x_{i+1} = (I-\tilde{G}F)x_i + \tilde{G}y, \end{cases}$$

and the *second* (the *dual*) *defect correction process*

$$\text{DCPB}\quad\begin{cases}\ell_0 = y, \quad x := \tilde{G}\ell_i, \\ \ell_{i+1} = (I-F\tilde{G})\ell_i + y.\end{cases}$$

REMARK 3. DCPA is completely described by $F,\tilde{G},y$ and x_0; DCPB is completely described by $F,\tilde{G},y$ and ℓ_0. With DCPA we use the fact that A is a linear space and not the fact that B is. With DCPB we use the fact that B is a linear space and not that A is.

REMARK 4. If $\tilde{G}$ is injective, then we can define its left-inverse $\tilde{F}$ and the DCPB can be shown to be equivalent with the iterative process

$$\text{DCPB}^*\quad\begin{cases}\tilde{F}x_0 = y, \\ \tilde{F}x_{i+1} = (\tilde{F}-F)x_i + y.\end{cases}$$

It is clear that, if $\hat{x}$ is a fixed point of the iteration DCPA then $\tilde{G}F\hat{x} = \tilde{G}y$. Hence, if $\tilde{G}$ is injective then $\hat{x}$ is a solution of the original problem (P). Also, if $\hat{\ell}$ is a fixed point of DCPB, then $F\tilde{G}\hat{\ell} = y$ and, hence, $\tilde{G}\hat{\ell}$ is a solution to (P).

If we consider the difference between the iterand x_i (resp. ℓ_i) and the fixed point $\hat{x}$ (resp. $\hat{\ell}$), then we notice that for linear F and $\tilde{G}$,

$$x_{i+1} - \hat{x} = (I-\tilde{G}F)(x_i-\hat{x}),$$

and

$$\ell_{i+1} - \hat{\ell} = (I-F\tilde{G})(\ell_i-\hat{\ell}).$$

Hence we call $M = I - \tilde{G}F$ the *amplification operator* (of the error) of DCPA and $\hat{M} = I - F\tilde{G}$ the amplification operator of DCPB. It is obvious that a sufficient condition for a DCP to converge to a fixed point is that the norm of its amplification operator is less than one. Generalizations for non-linear F and $\tilde{G}$ are obtained by local linearization, such as indicated in remark 5.5 below.

In Section 4 we shall need the following relation between $\hat{M}$ and M, which follows immediately from the definition

$$\hat{M} = FMF^{-1}.$$

THEOREM 1. *If* $\tilde{G}$ *is an affine mapping, then the sequences* $\{x_i\}$ *in* DCPA *and* $\{x_i\}$ *in* DCPB *are identical.*

PROOF. Let $\{\ell_i\}_{i=0,1,2,...}$ and $\{x_i\}_{i=0,1,2,...}$ be defined as in DCPB, then:

1) $x_0 = \tilde{G}\ell_0 = \tilde{G}y$, and

2) $x_{i+1} = \tilde{G}\ell_{i+1} = \tilde{G}(\ell_i - F\tilde{G}\ell_i + y) = \tilde{G}\ell_i - \tilde{G}F\tilde{G}\ell_i + \tilde{G}y$

$= x_i - \tilde{G}Fx_i + \tilde{G}y = (I - \tilde{G}F)x_i + \tilde{G}y;$

This means that the values from this sequence $\{x_i\}$ satisfy exactly the generation rules for the sequence $\{x_i\}$ from DCPA. Hence both sequences are identical. □

REMARK 5. It is clear from the proof of the last theorem that for a general mapping $\tilde{G}$ both processes DCPA and DCPB yield different sequences $\{x_i\}$.

A slight generalization of the DCPA, which is often more convenient for non-linear problems is the following defect correction process:

$$\text{DCPC}\quad \begin{cases} x_0 = \tilde{G}y \\ x_{i+1} = x_i + \mu\tilde{G}(\tilde{y} + (y - Fx_i)/\mu) - \mu\tilde{G}\tilde{y}. \end{cases}$$

In this iteration step the parameters μ and $\tilde{y}$ are still free to choose.

REMARKS. With respect to this new defect correction process we notice:

1. Near a solution of $Fx = y$ the operator $\tilde{G}$ is applied only in the neighbourhood of $\tilde{y}$.

In the general case (i.e. for any μ and $\tilde{y}$) a solution x_i of $Fx = y$ is a fixed point of DCPC.
With $\mu = -1$ and $\tilde{y} = y$, DCPC is identical with DCPA.
For arbitrary μ and $\tilde{y}$, with $\tilde{G}$ affine DCPC is identical with DCPA (and hence also with DCPB).
The amplification factor of DCPC is given by

$$\frac{\|x_{i+1}-\hat{x}\|}{\|x_i-\hat{x}\|} \le \|I-\tilde{G}'F'\| + \|\tilde{G}'\|\|F^*\| + \|\tilde{G}^*\|\|F'\| + \|\tilde{G}^*\|\|F^*\|,$$

ere $\tilde{G}'$ and $\tilde{G}^*$ are defined by

$$\tilde{G}(\tilde{y}+\delta)-\tilde{G}(\tilde{y}) = \tilde{G}'\delta + \tilde{G}^*\delta,$$

th $\tilde{G}'$ linear and $\tilde{G}^*$ such that

$$\|\tilde{G}^*\delta\| = o(\|\delta\|) \text{ as } \delta \to 0,$$

d F' and F^* defined analogously.

We conclude this section with some examples of defect correction ocesses.

AMPLE 1. Iterative methods for the solution of linear systems.

Many of the well-known iterative methods for the solution of near systems can easily be recognized as a defect correction process. r all these methods $\tilde{G}$ is linear and, hence, the three variants are uivalent. Here we shall identify as a DCP a number of these methods r the solution of the square linear system $Ax = b$.

.1. The Jacobi method

The Jacobi-method:

$$\text{diag}(A)\ x_{i+1} = b + (\text{diag}(A) - A)x_i,$$

an be written as

$$x_{i+1} = x_i + \tilde{G}(b-Ax_i) = (I-\tilde{G}A)x_i + \tilde{G}b,$$

with

$$\tilde{G} = [\mathrm{diag}(A)]^{-1}.$$

1.2. The Gauss-Seidel method

Let A be decomposed as $A = L + U$, where U is strictly upper-triangular and L is lower triangular; then the Gauss-Seidel process reads

$$Lx_{i+1} = b - Ux_i,$$

i.e. a DCP with $\tilde{G} = L^{-1}$.

1.3. The relaxation methods JOR, SOR, RF and GRF

All "stationary fully consistent iterative methods of degree one" for the solution of $Ax = b$ can be written as

$$x_{i+1} = x_i - P(Ax_i - b),$$

where P is a non-singular matrix (cf. YOUNG [1971]). With $P = pI$, p a scalar and I the identity matrix it is a stationary Richardson method (RF); with P a non-singular diagonal matrix it is a Generalized stationary Richardson method (GRF); with $P = \omega\tilde{G}$, $\tilde{G}$ as under 1.1 it is a Jacobi relaxation method (JOR) and with $P = \omega\tilde{G}$, $\tilde{G}$ as under 1.2 it is a SOR method.

EXAMPLE 2. Modified Newton iteration.

In this case the problem (P) is the solution of a non-linear equation

$$Fx = y,$$

with a Fréchet-differentiable operator F. The Fréchet-derivative $F'(x)$ is approximated by a non-singular linear operator E. The relation

$$Fx - Fx_i = F'(x_i)(x-x_i) + o(\|x-x_i\|),$$

or equivalently,

$$x - x_i = (F'(x_i))^{-1}(y-Fx_i+o(\|x-x_i\|)),$$

suggests the modified Newton iteration:

$$x_{i+1} = x_i + E^{-1}(y-Fx_i).$$

Clearly, this is a DCPA with $\tilde{G} = E^{-1}$.

We notice that in a proper Newton process (not the modified Newton iteration) the approximate Fréchet-derivative E is updated during the iteration process. This kind of generalization of the elementary DCP will be treated in Section 2.

EXAMPLE 3. An analytic example.

We consider the two-point boundary-value problem (cf. STETTER [1978])

$$(*) \qquad \begin{cases} x'' - e^x = 0 & \text{on } (-1,+1) \\ x(-1) = x(+1) = 0. \end{cases}$$

This defines the problem

$$Fx = 0,$$

where

$$F: C_0^2[-1,+1] \to C(-1,+1).$$

We construct an approximate problem, replacing e^x by $0.99 + 0.81x$ (i.e. a reasonable approximation if $-0.4 \le x \le 0.0$). Thus we get the approximate problem $\tilde{F}x = y$, viz.

$$\begin{cases} x'' - 0.81x - 0.99 = y & \text{on } (-1,1) \\ x(-1) = x(+1) = 0. \end{cases}$$

Hence, we can write the solution of $\tilde{F}x = y$ as

$$x(t) = \int_{-1}^{+1} K(t,z)(y(z) + 0.99) \, dz,$$

for some suitable kernel-function $K(t,z)$. This integral operator defines an approximate inverse $\tilde{G}$ for the problem (*). With this $\tilde{G}$ we can construct a DCP to find the solution of (*).

2. EXTENSION OF THE DCP PRINCIPLE

In this section we shall extend the idea of the defect correction process in several ways: we allow different approximate inverses to serve in one iteration process and we consider a sequence of problems that converges to a final problem of which the solution is wanted. We also consider the process obtained when a fixed combination of approximate inverses is used all over in a defect correction process.

2.1. Non-stationary defect correction processes

In order to find a solution to the problem (P) it is not necessary to use one fixed approximate inverse in an iteration process as described in the preceeding section. As we anticipated in the example with Newton's method, it is possible to use another approximate inverse in each iteration step. Then the iteration steps in DCPA and DCPB read respectively

$$x_{i+1} = x_i - \tilde{G}_i F x_i + \tilde{G}_i y$$

and

$$\ell_{i+1} = \ell_i - F\tilde{G}_i \ell_i + y.$$

A similar modification for DCPC can be given.

Various methods are known to find a proper sequence of $\{\tilde{G}_i\}$.

Here we mention a few.

EXAMPLE 1. $\tilde{G}_i = \tilde{G}(x_{i-1})$.

The approximate inverse depends on the last iterand computed. This is the case e.g. in Newton's method for the solution of non-linear equations, where $\tilde{G}(x) = (F'(x))^{-1}$; $F'(x)$ is the Fréchet derivative of the operator in the problem (P).

EXAMPLE 2. $\tilde{G}_i = \tilde{G}(\omega_i)$.

The approximate inverse depends on a single real parameter. This is the case e.g. in non-stationary relaxation processes for the solution of linear systems.

EXAMPLE 3. $\tilde{G}_i \in \{\tilde{G}_1, \tilde{G}_2\}$.

In each iteration step the approximate inverse is chosen out of a set of two (or more) fixed approximate inverses. This is the case e.g. in Brakhage's and Atkinson's methods for the solution of Fredholm integral equations of the 2nd kind. (See ATKINSON [1976] and BRAKHAGE [1960].)

2.2. A fixed combination of approximate inverses

We consider two iteration steps in the non-stationary DCPA in which, in turn, one or the other of two approximate inverses is used. In the linear case, the iteration steps

$$x_{i+\frac{1}{2}} = (I-\tilde{G}F(x_i + \tilde{G}y$$
$$x_{i+1} = (I-\tilde{\tilde{G}}F(x_{i+\frac{1}{2}} + \tilde{\tilde{G}}y$$

combine into a single iteration step of the form

$$x_{i+1} = (I-\tilde{\tilde{G}}F)(I-GF)x_i + (\tilde{\tilde{G}}-\tilde{\tilde{G}}FG+\tilde{G})y.$$

This is easily recognized as a new iteration step of the type DCPA, now with the approximate inverse

$$\hat{G} = \approx{G} - \approx{G}F\tilde{G} + \tilde{G}.$$

We conclude that a fixed combination of DCPA-steps can be considered as a new DCPA-step with a more complex approximate inverse. The amplification operator of the new DCPA process is the product of the amplification operators of the elementary processes.

REMARK. Generally the above observation with respect to DCPA does not directly hold for DCPB processes.

2.3. σ applications of the same approximate inverse

In order not to make the notation unnecessarily intricate, from now on we shall only consider linear problems, unless explicitly stated otherwise.

We can describe the DCPA in matrix notation by

$$\begin{pmatrix} x_{i+1} \\ y \end{pmatrix} = \begin{pmatrix} I-\tilde{G}F & \tilde{G} \\ \emptyset & I \end{pmatrix} \begin{pmatrix} x_i \\ y \end{pmatrix} .$$

σ times an application of the same iteration step yields

$$\begin{pmatrix} x_{i+\sigma} \\ y \end{pmatrix} = \begin{pmatrix} I-\tilde{G}F & \tilde{G} \\ \emptyset & I \end{pmatrix}^{\sigma} \begin{pmatrix} x_i \\ y \end{pmatrix} = \begin{pmatrix} (I-\tilde{G}F)^{\sigma} & \sum_{m=0}^{\sigma-1} (I-\tilde{G}F)^m \tilde{G} \\ \emptyset & I \end{pmatrix} \begin{pmatrix} x_i \\ y \end{pmatrix} .$$

Thus, one iteration step which consists of σ applications of DCPA-steps results in a DCPA with the amplification operator

$$M = (I-\tilde{G}F)^{\sigma}$$

and the approximate inverse

$$\hat{G} = \sum_{m=0}^{\sigma-1} (I-\tilde{G}F)^m \tilde{G} = [I - (I-\tilde{G}F)^{\sigma}]F^{-1}.$$

2.4. Iterative application of DCP

It is possible not only to change the approximate inverse $\tilde{G}$ during the iteration process, often it makes sense also to substitute different operators F_i for F during iteration. In general, the operators $\{F_i\}$ will be simple to evaluate in the beginning of the iteration and they will converge to F, the operator in the original problem, as the iteration proceeds.

One example of such a process is the IUDeC (Iteratively Updated Defect Correction) process described by STETTER [1978]. Here $\{F_i\}$ are discrete approximations of higher and higher order to an analytic operator F. The approximate inverse $\tilde{G} = F_0^{-1}$ is kept constant during the process. An analysis of this kind of process is given in Section 3.3, when we have introduced discretizations.

Another example is the Full Multigrid method [BRANDT, 1979] in which $\{F_i\}$ are discretizations on finer and finer nets of an analytic operator F.

2.5. Recursive application of DCP

Generally, the evaluation of the approximate inverse operator $\tilde{G}_i$ implies the solution of an equation which is (essentially) of a simpler type than the original equation. However, also this simpler equation may be of a kind that we want to solve by means of a DCP. For this we need an even simpler equation to solve, etc.. Thus, the execution of a single iteration step may imply the activation of a new (simpler to solve) DCP.In this way we can construct a recursive construction of DCPs in which only on the lowest level of recursion a very simple equation is to be solved.

Independently, this is probably not a real meaningful construction, but in combination with non-stationary processes, where also other (non-recursive) approximate inverses are available, it describes the essentials of the multigrid algorithm.

Such a combination of a non-stationary process with some recursive approximate inverses can be described by the following sequence of DCPs.

$$
\begin{array}{llll}
DCP_1: & x := x - \tilde{G}_1\,(F_1x - f_1) & \tilde{G}_j \text{ fixed}, & j=1,2,\ldots,n\,, \\
DCP_2: & x := x - \tilde{G}_{2,i}(F_2x - f_2) & & \\
\vdots & \vdots & \tilde{G}_{j,i} \in \{\tilde{G}_j,\, F_{j-1}^{-1}\}, & \\
 & & & j=2,3,\ldots,n. \\
DCP_n: & x := x - \tilde{G}_{n,i}(F_nx - f_n) & &
\end{array}
$$

A full use of the sequence of DCPs is made by combining also the iterative application: first DCP_1 is solved and its solution is used as a starting value for DCP_2 etc.. In a multigrid context

$$DCP_1, DCP_2, \ldots, DCP_n,$$

are processes to solve operator equations, discretized on finer and finer grids. The complete iterative process is called: Full Multigrid Algorithm [BRANDT, 1979].

3. DISCRETIZATION ON RELATED GRIDS, RELATED DISCRETIZATIONS

In this section we give definitions for related discretizations of spaces and problems and we define relative order of approximation, consistency and convergence between related discretizations. In Section 3.3 we give an approximation theorem for successive approximations in the iterative application of a DCP.

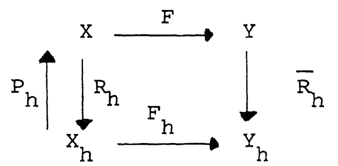
3.1. Discretization of spaces and operators

Let's be given a problem $Fx = y$, where $F:X \to Y$ and $y \in Y$ are given and where X and Y are (infinite dimensional) vector spaces. The problem is discretized by associating it with a problem $F_h\, x_h = y_h$, where $F_h: X_h \to Y_h$ and $y_h \in Y_h$ are given and X_h and Y_h are finite dimensional vector spaces. By selecting $h \in \mathcal{H}$ ($\mathcal{H}$ an index-set) different discretizations of the same problem are possible.

A relation between the problem and its discretization is obtained by introducing surjections $R_h: X \to X_h$ and $\bar{R}_h: Y \to Y_h$. (Notice that $\dim(X) \geq \dim(X_h)$, $\dim(Y) \geq \dim(Y_h)$ and, in most cases, X_h and Y_h are selected such that $\dim(X_h) = \dim(Y_h)$.)

In order to interpret the solution of the discretized problem as an approximation to the solution of the original problem, we have to define an injection $P_h : X_h \to X$.

The mappings P_h are called *prolongations*, the mappings R_h and $\bar{R}_h$ are *restrictions*. The relation between the different spaces and mappings is summarized in the following diagram



$h \in H$

DEFINITION. Given the discretization of the spaces X and Y by X_h, Y_h, P_h, R_h and $\bar{R}_h$, $h \in H$, we can associate with the problem $Fx = y$ its *Galerkin discretization* $F_h x_h = y_h$ by defining $F_h = \bar{R}_h F P_h$ and $y_h = \bar{R}_h y$.

DEFINITION. Given to discretizations of the spaces X and Y by $(X_h, Y_h, P_h, R_h, \bar{R}_h)$ and $(X_H, Y_H, P_H, R_H, \bar{R}_H)$, $h, H \in H$, these are called *related discretizations* if surjective mappings R_{Hh} and $\bar{R}_{Hh}$ and an injection P_{hH} exist such that

$$\begin{aligned} R_{Hh} &: X_h \to X_H, & R_{Hh} R_h &= R_H, \\ \bar{R}_{Hh} &: Y_h \to Y_H, & \bar{R}_{Hh} \bar{R}_h &= \bar{R}_H, \\ P_{hH} &: X_H \to X_h, & P_h P_{hH} &= P_H. \end{aligned}$$

It should be clear that $\dim(X_H) \le \dim(X_h)$ and $\dim(Y_H) \le \dim(Y_h)$. We see also that, if two discretizations (with $h, H \in H$) of the spaces X and Y are related, then the *coarse discretization* (with $H \in H$) can be considered as a discretization of the *fine discretization* (with $h \in H$) of the finite dimensional spaces X_h and Y_h.

From the definitions it follows immediately that, if the coarse discretization $F_H x_H = Y_H$ and the fine discretization $F_h x_h = y_h$ are both Galerkin discretizations of the same problem $Fx = y$, we have

$$F_H = \bar{R}_{Hh} F_h P_{hH} \quad \text{and} \quad y_H = \bar{R}_{Hh} y_h.$$

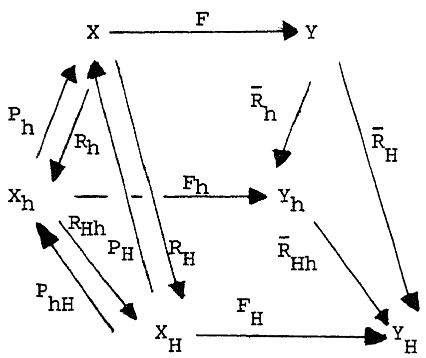
Because P_h is an injection, it has a left-inverse $\hat{R}_h$ such that $\hat{R}_h P_h$ is the identity operator on X_h; because R_h and $\bar{R}_h$ are surjective, right-inverses $\hat{P}_h$ and $\bar{P}_{\hat{h}}$ exist such that $R_h \hat{P}_h : X_h \to X_h$ and $\bar{R}_h \bar{P}_{\hat{h}} : Y_h \to Y_h$ are identity operators. From these definitions of $\hat{R}_h$, $\hat{P}_h$ and $\bar{P}_{\hat{h}}$ follows:

$$R_{Hh} = R_H \hat{P}_h ,$$

$$\bar{R}_{Hh} = \bar{R}_H \bar{P}_{\hat{h}} ,$$

$$P_{hH} = \hat{R}_h P_H .$$

The relation between the different spaces is summarized in the following diagram.



It is important to notice that, in general, different norms can be used to (trans-) form the above mentioned vector spaces into normed vector spaces or Banach spaces. Indeed, each of the above vector spaces, say Z, can be formed into a scale of normed vector spaces $\{Z^\alpha\}$, $\alpha \in \mathbb{R}$, with $Z^\alpha = Z$ and norms $\|\cdot\|_\alpha$ such that with $u \in Z$ we have

$$\|u\|_\alpha \le \|u\|_\beta \text{ if } \alpha \le \beta.$$

<u>DEFINITIONS</u>. An operator $F : X \to Y$ is called *bounded* if

$$\|F\|_{X^\alpha \to Y^\alpha} \le C \qquad \text{uniformly in } \alpha,$$

and σ-*stable* if

$$\|F^{-1}\|_{Y^\alpha \to X^{\alpha-\sigma}} \le C \qquad \text{uniformly in } \alpha,$$

In the following we shall assume that all restrictions and prolongations and their right-resp. left-inverses are bounded, uniformly in $h \in H$. The conditions on the inverses imply for the prolongations P_h that

$$\inf_{v \neq 0} \frac{\|P_h v\|}{\|v\|} > C > 0, \qquad C \text{ independent of } h \in H,$$

and for the restrictions R_h that

$$\inf_{\substack{w \in R_h \\ w \neq 0}} \sup_{\{v | R_h v = w\}} \frac{\|R_h v\|}{\|v\|} > C > 0, \quad C \text{ independent of } h \in H.$$

To each discretization, characterized by $h \in H$, a mesh-size $m(h) > 0$ is associated. Discretizations X_h and X_H of X with $\dim(X_h) \ge \dim(X_H)$ generally have mesh-sizes related by $m(h) < m(H)$. If no confusion is possible we denote $m(h)$ simply by h. Often we consider infinite sequences $\{X_h\}$ with $h > 0$ and $\lim_{h\to 0} \dim(X_h) = \infty$.

3.2. Relative consistency and convergence

DEFINITIONS. A sequence of discretizations of X characterized by $(X_h, P_h, R_h)_{h>0}$ is called *convergent* if

$$\lim_{h\to 0} \|I - P_h R_h\| = 0;$$

the *order of approximation* is p if

$$\|I - P_h R_h\| = O(h^p) \text{ for } h \to 0.$$

DEFINITION. A sequence of discretizations of a problem $Fx = y$ is *consistent* if

$$\lim_{h\to 0} \|F_h R_h - \bar{R}_h F\| = 0;$$

its *order of consistency* is p if

$$\|F_H R_h - \bar{R}_h F\| = \mathcal{O}(h^p) \quad \text{for } h \to 0.$$

<u>DEFINITION</u>. A sequence of discretizations of a problem $Fx = y$ is *σ-stable* if $F_h^{-1}: Y_h^\alpha \to X_h^{\alpha-\sigma}$ is bounded uniformly in h and α. It is called *stable* if it is 0-stable.

<u>DEFINITION</u>. A sequence of discretizations of a problem $Fx = y$ is *convergent* if

$$\lim_{h\to 0} \|F^{-1} - P_h F_h^{-1} \bar{R}_h\| = 0$$

its *order of convergence* is p if

$$\|F^{-1} - P_h F_h^{-1} \bar{R}_h\| = \mathcal{O}(h^p) \quad \text{for } h \to 0.$$

Analogously, for related discretizations characterized by $H > h > 0$, we can define the corresponding relative properties (without reference to the original problem), i.e.

the *relative order of approximation* p:

$$\|I_h - P_{hH} R_{Hh}\| = \mathcal{O}(H^p),$$

the *relative order of consistency* p

$$\|F_H R_{Hh} - \bar{R}_{Hh} F_h\| = \mathcal{O}(H^p),$$

the *relative order of convergence* p:

$$\|F_h^{-1} - P_{hH} F_H^{-1} \bar{R}_{Hh}\| = \mathcal{O}(H^p).$$

<u>THEOREM</u>. *If two related discretizations of the same problem are consistent of order* p_1 *and* p_2 *respectively, then they are relatively consistent of the order* $\min(p_1, p_2)$.

<u>PROOF</u>. The simple proof is left to the reader.

NOTE 1. The following identity is useful if we consider DCPs with related discretizations

$$I_h - P_{hH}F_H^{-1}\bar{R}_{Hh}F_h = (I_h - P_{hH}R_{Hh}) + P_{hH}F_H^{-1}(F_H R_{Hh} - \bar{R}_{Hh}F_h).$$

NOTE 2. Let $F_h x_h = y_h$ and $F_H X_H = Y_H$ be two related Galerkin discretizations of the same problem, then, for any restriction $\tilde{R}_{Hh} : X_h \to X_H$ we have

$$I_h - P_{hH}F_H^{-1}\bar{R}_{Hh}F_h = (I_h - P_{hH}F_H^{-1}\bar{R}_{Hh}F_h)(I_h - P_{hH}\tilde{R}_{Hh}).$$

3.3. The accuracy of successive approximations in a DCP iteration with different discretizations on the same problem

Let us consider (different) discretizations of the problem $Fx = y$, viz.

$$F_h^i x_h = y_h, \quad \text{with } F_h^i : X_h \to Y_h \quad \text{for all } i = 0,1,2,\ldots,$$

and let X, X_h, Y and Y_h be related by

$$R_h : X \to X_h \quad \text{and} \quad \bar{R}_h : Y \to Y_h.$$

Let the order of consistency of the discretizations be p_i, and let the first discretization be stable. We will study the iterative application of DCPA, with the equations $F_h^i x_h = y_h = \bar{R}_h y$ to solve in the i-th iteration step and with the same approximate inverse $\tilde{G}_h = (F_h^0)^{-1}$ in all iteration steps. Then the DCPA reads

$$\begin{cases} u_1 = \tilde{G}_h y_h = \tilde{G}_h \bar{R}_h y \\ u_{i+1} = (I_h - \tilde{G}_h F_h^i)\, u_i + \tilde{G}_h y_h. \end{cases}$$

We are going to estimate the relative error of approximation for a finite number of iteration steps:

$$k_i = \| u_i - R_h x \| / \| x \|.$$

THEOREM. *For the relative error of approximation in the i-th iteration step of the iterative DCPA process:*

$$k_i = \|u_i - R_h x\| / \|x\|,$$

we have

$$k_0 = \|\tilde{G}_h\| \, \|\bar{R}_h F - F_h^0 R_h\| = O(h^{p_0})$$

$$k_i = \|\tilde{G}_h\| \, \|\bar{R}_h F - F_h^{i-1} R_h\| + \|\tilde{G}_h\| \, \|F_h^1 - F_h^{i-1}\| \, k_{i-1}$$

$$= O(h^{\min_{0\le j\le i} (p_j + (i-j)p_0)}), \quad i = 1,2,\dots \quad .$$

PROOF.

$$u_0 - R_h x = \tilde{G}_h \bar{R}_h \, y - R_h \, x = \tilde{G}_h (\bar{R}_h F - F_h^0 \, R_h) \, x.$$

The given estimate now follows from the stability of F_h^0 (i.e. $\tilde{G}_h$ is uniformly bounded) and the consistency of F_h^0.

$$u_{i+1} - R_h \, x = u_i - R_h x - \tilde{G}_h F_h^i \, u_i + \tilde{G}_h \, y_h$$

$$= u_i - R_h x + \tilde{G}_h (\bar{R}_h F_x - F_h^i R_h x + F_h^i R_h x - F_h^i u_i)$$

$$= (I_h - \tilde{G}_h F_h^i)(u_i - R_h x) + \tilde{G}_h (\bar{R}_h F - F_h^i R_h) x \quad .$$

Hence, for i = 0,1,2,...,

$$k_{i+1} \le \| I_h - G_h F_h^i \| \, k_i + \|\tilde{G}_h\| \, \|\bar{R}_h F - F_h^i \, R_h\|.$$

Here again, the estimate follows from the stability of F_h^0 and the consistency of F_h^i. □

COROLLARY. *If*

$$\begin{cases} p_i \geq (i+1) & (0 \leq i < n), \\ p_i = p_n & (i \geq n) \end{cases}$$

then

$$k_i = \mathcal{O}(h^{\min(p_n,(i+1)p_0)}).$$

4. MULTIGRID ALGORITHMS

In this section we shall describe multigrid algorithms and the structure of their convergence theorems. First we consider a simple form of the multigrid algorithm, "the two-level algorithm" (or TLA), and show how its convergence is proved. Then we show the multi-level algorithm (MLA), which is the recursive application of the two-level algorithm. At the end we show how multigrid algorithms are applied to non-linear problems.

The problems that are solved by multigrid methods are all discretizations of a continuous problem $Lx = f$. The methods find solutions to the finest discretization $L_h x_h = f_h$ by means of discretizations on coarser grids, which we denote by $L_H x_H = f_H$.

4.1. The two-level algorithm

The two-level algorithm is a non-stationary defect correction process in which only two different approximate inverses are used:

(1) some *relaxation method* (e.g. Jacobi, Gauss-Seidel or the incomplete LU-decomposition, see example 1 Section 1) on the fine grid and

(2) a *coarse grid correction*.

The approximate inverse in the coarse grid correction that is used to solve the discrete problem $L_h x_h = f_h$ is given by $\tilde{G}_i = P_{hH} L_H^{-1} \bar{R}_{Hh}$. Thus, one coarse grid correction step in the two-level algorithm reads

$$x_{i+1} = x_i + P_{hH} L_H^{-1} \bar{R}_{Hh} (f_h - L_h x_i).$$

One step in the two-level algorithm, now consists of p relaxation sweeps of the relaxation method chosen, a coarse grid correction step and again q relaxation sweeps of the relaxation method. Such a step of the two-level algorithm is described in the following ALGOL-like program:

```
proc two level algorithm = (ref gridf u, gridf f) void:
begin
      for i to p
      do relax (u,f) od;

      d := restrict (Lh u - f);
      solve (v,d);                      # solves L_H v = d #
      u := u - prolongate v ;

      for i to q
      do relax (u,f) od
end;
```

Clearly, the amplification operator of one step of the two-level algorithm is given by

$$M_h^{TLA} = (I - B_h L_h)^q \; (I - P_{hH} \, L_H^{-1} \, \bar{R}_{Hh} \, L_h)(I - B_h \, L_h)^P,$$

where B_h is the approximate inverse of the relaxation process. In this expression we recognize the relative convergence operator and the amplification operators of the relaxation process:

$$M_h^{REL} = (I - B_h \, L_h),$$
$$\hat{M}_h^{REL} = (I - L_h \, B_h),$$

and we can write

$$M_h^{TLA} = (M_h^{REL})^q \; (L_h^{-1} - P_{hH} \, L_H^{-1} \, R_{Hh})(\hat{M}_h^{REL})^P \; L_h,$$

or

$$\hat{M}_h^{TLA} = L_h \ (M_h^{REL})^q \ (L_h^{-1} - P_{hH} L_H^{-1} R_{Hh}) \ (\hat{M}_h^{REL})^p .$$

The structure of the convergence proof for the two-level algorithm is as follows:

Assuming that

(1) the two discrete operators are relatively convergent of order α,

(2) the relaxation satisfies a *proper smoothing property* of order at least α, i.e. $\exists\ C_0(p) > 0$, independent of h, such that $\| (\hat{M}_h^{REL})^p L_h \| < C_0(p) h^{-\alpha}$ and $\lim_{p\to\infty} C_0(p) = 0$,

(3) the amplification operator $(M_h^{REL})^q$ is bounded,

(4) the mesh-ratio $m(H)/m(h)$ is bounded, uniformly in h, then the two-level algorithm converges for p large enough.

<u>PROOF</u>.

$$\begin{aligned} \|M_h^{TLA}\| &\le \|(M_h^{REL})^q\| \ \|L_h^{-1} - P_{hH} L_H^{-1} \bar{R}_{Hh}\| \ \|(\hat{\Gamma}_h^{REL})^p L_h\| \\ &\le C \ . \ C \ (m(H))^{\alpha} \ . \ C_0(p) \ (m(h))^{-\alpha} \\ &= C \ . \ C_0(p) \ (m(H)/m(h))^{\alpha} \le C \ . \ C_0(p) \end{aligned}$$

Since $C_0(p) \to 0$ for $p \to \infty$ we see that $\|M_h^{TLA}\| < 1$ for p large enough. □

<u>REMARK</u>. In an actual convergence proof the norms in the relevant spaces should be specified and the assumptions should be verified for the particular algorithm under consideration. We have to realize that, apart from the above mentioned structure, the two-level algorithm is determined by the particular discretizations L_h and L_H, by the restrictions and prolongations $\bar{R}_{Hh}$ and P_{hH} and by the particular relaxation method used (characterized by B_h).

If the discretizations L_h and L_H are related Galerkin discretizations, then we can make use of the relations in the notes 1 and 2 of Section 3.2.

4.2. <u>The multi-level algorithm</u>

Whereas for the two-level algorithm we have to evaluate L_H^{-1}, i.e.

we have to solve a discretized problem on a coarse grid, in the multi-level algorithm we approximate this solution by application of a number of iteration steps of the same algorithm on the coarse level. As was explained in Section 2.5 we now only have to solve directly a discretized problem on the very coarsest grid. If σ iteration steps of the multi-level algorithm are used to approximate L_H^{-1}, this multi-level algorithm is described in the ALGOL-like program:

```
proc multi level algorithm = (ref gridf u, grdif f) void:
begin
      for i to p while ...
      do relax (u,f) od;

      d := restrict (f - Lh u); v := 0;
      if level of u = 1
      then solve (u,f)
           # on the coarsest grid #
      else for i to sigma while ...
           do multi level algorithm (u,d) od
      fi;
      u := u + prolongate v

      for i to q while ...
      do relax (u,f) od
end;
```

By while ... we denote in the program that some iterations may be terminated sooner, depending on the speed of convergence or other conditions that can be checked during the computation. Multigrid algorithms that make use of this possibility are said to have an *adaptive strategy*, algorithms where the iterations are controled only by the fixed numbers p, σ and q are said to have a *fixed strategy*. Although the adaptive strategy may be very efficient (cf. BRANDT, 1979), the fixed strategy is better accessible for a theoretical analysis.

For some fixed strategies, we show in figure 1 how it is switched between the different levels of discretization. We see that - essentially - most relaxation sweeps are performed on the lower levels.

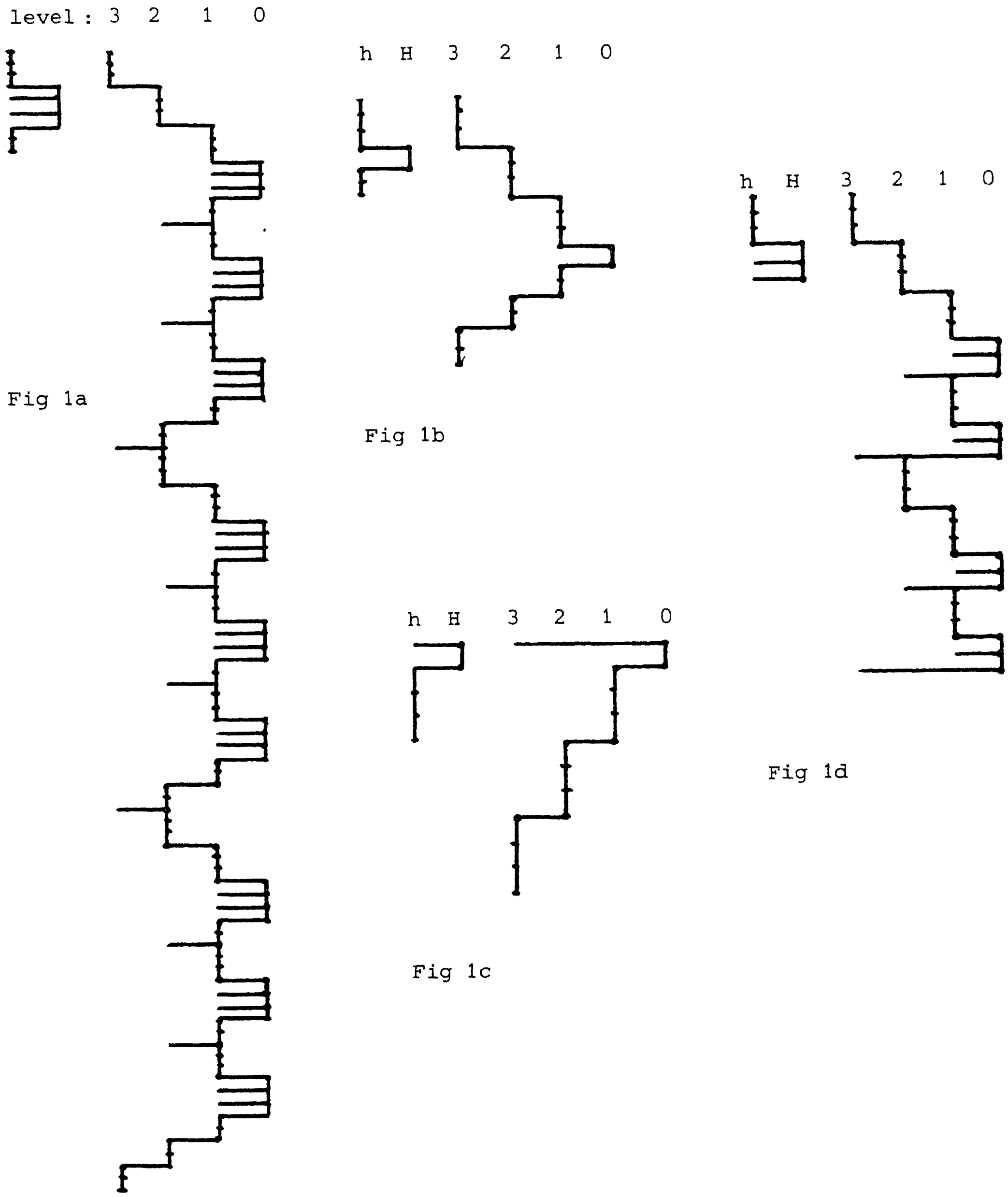


Figure 1. The recursive structure of multigrid algorithms with a fixed strategy.

In all diagrams the number of levels is 3, the very coarsest level is denoted by 0. In each diagram 1a, 1b, 1c or 1d, the basic structure on the levels h and H is given as well as the recursive structure of one iteration step on level 3. Tick marks on a level > 0 denote the execution of a relaxation step on this level, tick-marks on level 0 denotes the direct solution on the very coarsest level. The different structures shown are:

1a. A general structure with $p = 3$, $\sigma = 3$, and $q = 2$.
1b. A structure with $\sigma = 1$ (NICOLAIDES, 1979) $p = 3$, $q = 2$.
1c. A structure with $\sigma = 1$, $p = 0$ (FREDERICKSON, 1975) $q = 3$.
1d. A structure with $q = 0$ (HACKBUSCH, 1979) $p = 3$, $\sigma = 2$.

The amplification operator of a multi-level iteration step on the h-level of discretisation we denote by M_h^{MLA}, this amplification operator on the next coarser level we denote by M_H^{MLA}. The approximate inverse of the coarse grid correction in the multigrid algorithm is not given by L_H^{-1}, but it is obtained by σ steps in the DCP for the approximation of L_H^{-1}. The amplification operator of such a single DCP-step is given by M_H^{MLA}. Hence, the approximate inverse of the σ iteration steps together is given by (see Section 2.3):

$$(I - (M_H^{MLA})^\sigma)L_H^{-1}.$$

Consequently, the amplification operator of the coarse grid correction is

$$(I - P_{hH}(I - (M_H^{MLA})^\sigma)\ L_H^{-1}\ \bar{R}_{Hh}\ L_h)$$

and we have

$$M_h^{MLA} = (M_h^{REL})^q (I - P_{hH}(I - (M_H^{MLA})^\sigma)L_H^{-1}\ \bar{R}_{Hh}\ L_h)\,(M_h^{REL})^p$$

$$= M_h^{TLA} + (M_h^{REL})^q\ P_{hH}(M_H^{MLA})^\sigma\ L_H^{-1}\ \bar{R}_{Hh}(\hat{M}_h^{REL})^p L_h$$

or

$$\hat{M}_h^{MLA} = \hat{M}_h^{TLA} + L_h(M_h^{REL})^q\ P_{hH}(M_H^{MLA})^\sigma L_H^{-1}\ \bar{R}_{Hh}(\hat{M}_h^{REL})^p.$$

Therefore, if the (coarse) discretized operator L_H is stable and the assumptions (2) and (3) of Section 4.1 hold, then

$$\|M_h^{MLA}\| \le \|M_h^{TLA}\| + C\|M_H^{MLA}\|^{\sigma}.$$

Here we get a recursive expression, where the rate of convergence of the MLA on the level h is expressed in the rate of convergence of the TLA and the rate of convergence of the MLA on the next coarser level H. Further we notice that on the coarsest level we have $M_0^{MLA} = M_0^{TLA}$.

On each level we have $\|M_h^{TLA}\| \le \rho < 1$ if p is large enough, hence we can find a σ such that $\|M_H^{MLA}\| < 1$. Often a small value of σ (e.g. σ=2) can be shown to be sufficient to have $\|M_h^{MLA}\| \le \rho < 1$ on all levels, ρ independent of h.

4.3. The non-linear multi-level algorithm

The multi-level algorithm in Section 4.2 essentailly used the fact that the operator L and its discretizations are linear. By a slight change of the algorithm we can adapt it for nonlinear problems. For this purpose we make use of the DCPC as treated in Section 1. We describe the nonlinear algorithm - again - in an ALGOL-like program

```
proc non linear mla = (ref gridf u, gridf f) void:
if level of u = 0
then solve (u,f)
        # e.g. by a Newton type method #
begin
        for i to p
        do relax (u,f) od;

        y := w := restrict u;
        d := LH y + restrictbar (f - Lh u)/mu;
        if level of u = 1
        then solve (u,f)
            # e.g. by some Newton methods #
        else for i to sigma
```

```
            do nonlinear mla(w,d) od
        fi;
        u := u + mu × prolongate (w-y);

        for i to q
        do relax (u,f) od
    end;
```

Here, of course, the relaxation should be of a non-linear type. The coarse grid correction of the TLA corresponding with this MLA (i.e. the MLA with $\sigma = \infty$) is here

$$x_{i+1} = x_i + \mu P_{hH}(L_H^{-1}(L_H R_{Hh} x_i + \bar{R}_{Hh}(f_h - L_h x_i)/\mu) - R_{Hh} x_i).$$

This can be recognized as the DCPC in Section 1, with $\tilde{y}$ such that $L_H R_{Hh} x_i = R_{Hh} \tilde{y}$.

If we fit the nonlinear MLA-step into a Full Multigrid Method (see Section 2.4), then we may replace $R_{Hh} x_i$ (i.e. the best approximation of the solution that is available at the level H) by the last solution obtained on the next coarser grid. In that case, there is no need for recomputing y and LH y in each call of the nonlinear MLA.

5. EXAMPLES OF MULTIGRID METHODS

In this section we give two examples of multigrid methods. In the first example we show Fredericson's method for the solution of a differential equation and in the second we treat a multigrid method for the solution of a Fredholm integral equation of the 2nd kind. The essential difference between both problems is that a regular differential operator, $L : A \to B$, maps a space with a stronger into a space with a weaker topology, whereas a compact integral operator, $K : A \to B$, maps a space with a weaker into one with a stronger topology. The effect is, that for the differential equation we can get an amplification factor $\|M_h^{MLA}\|$ which is bounded by a constant (less than one) uniformly in h. We call this a *multigrid method of the first kind*. For the integral equation we can get an amplification factor $\|M_h^{MLA}\|$ which is bounded by a constant of order $\mathcal{O}(h^m)$ for some $m > 0$. This we call a *multigrid method of the second kind*.

REMARK. With Jacobi-type iteration similar differences are found for the two different problems: for the differential equation we have the bound $\|M_h^{REL}\| \le 1 - Ch^{2m}$ and for the integral equation the bound is $\|M_h^{REL}\| \le C < 1$ as $h \to 0$. These bounds also clearly show the supremacy of the MLA-iteration over the classical iteration methods.

5.1. The multigrid method of Fredericson for the solution of a differential equation

For Fredericson's multigrid method we have $p = 0$ and $\sigma = 1$. Because of $\sigma = 1$ the amplification operator is much simpler than in the general case. For a 3-level method (see figure 1.c) this operator is given by

$$\begin{aligned} M_3^{MLA} = {} & (I-B_3L_3)^q (L_3^{-1}-PL_2^{-1}R)\ L_3 \\ & + (I-B_3L_3)^q\ P(I-B_2L_2)^q (L_2^{-1}-PL_1^{-1}R)R\ L_3 \\ & + (I-B_3L_3)^q\ P(I-B_2L_2)^q\ P(I-B_1L_1)^q (L_1^{-1}-PL_0^{-1}R)RRL_3, \end{aligned}$$

where L_i is the discretized operator at level i, $(I-B_iL_i)$ is the amplification operator of the relaxation at level i, and P and R are the prolongation and restriction operators between the various levels. First we look at the first term of this operator:

$$(I-B_3L_3)^q(I-PL_2^{-1}RL_3).$$

Here $(I-PL_2^{-1}RL_3)$ reduces the low frequences in the error and $(I-B_3L_3)^q$ reduces the high frequencies in the error of the approximation to the solution. This can be seen e.g. if L_2 and L_3 are related canonical discretizations: $L_2 = RL_3P$. Then the first term can be rewritten as

$$(I-B_3L_3)^q(I-PL_2^{-1}RL_3)(I-P\tilde{R}).$$

If $\tilde{R}$ denotes restriction to gridponts and P denotes piecewise

polynomial interpolation of degree k-1 then it is clear that for $I - P\tilde{R} : H^k \to H^0$ we have $\| I-P\tilde{R}\|_{H^k \to H^0} \le Ch^k$.
$(I-PL_2^{-1}RL_3): H^0 \to H^0$ being bounded we need for smoothing property

$$\| (I-B_3L_3)^q\|_{H^0 \to H^k} \le C(q)h^{-k}$$

with C(q) sufficiently small for large enough q, i.e. components in the error with large derivatives should be damped sufficiently. Such estimates can be proved. E.g. HACKBUSCH [1979] proves for regular elliptic differential problems of order 2m and (damped) Jacobi relaxation:

$$\| (I-B_3L_3)^q\|_{H^\alpha \to H^{\alpha+2m}} \le q^{-1}\ h^{-2m}.$$

Analogously, in the third term of M_3^{MLA}, the factor $L_1^{-1} - PL_0^{-1}R$ reduces the lowest frequencies, whereas the factors $(I-B_iL_i)$, = i,1,2,3, reduce each a particular range of higher frequencies. The final effect is that a bound for $\|M_h^{MLA}\|$ can be found that is less than one *uniformly in h*. This is in contrast with a plain relaxation method for the solution of a discretized differential equation for which $\|M_h^{REL}\| \to 1$ as $h \to 0$.

5.2. A multigrid method for the solution of a Fredholm integral equation of the 2nd kind

In this example we consider the integral equation

$$x(s) - \int_a^b k(s,t)x(t)\ dt = y(s),$$

or, in operator notation,

$$Lx \equiv x - Kx = y,$$

and we consider a sequence of related discretizations

$$L_p\, x \equiv x - K_p x = y_p, \qquad p = 0,1,2,\ldots ,$$

with $h_p \to 0$ as $p \to \infty$.
A simple method to solve the discrete equation is by means of successive substitution

$$x_{i+1} = K_p x_i + y_p .$$

This is a Jacobi-type iteration: it is a DCPA with approximate inverse $\tilde{G} = I$. It converges if $\|K_p\| < 1$ and, for a compact operator K, it has a smoothing property.

For $p > 0$, also a coarse grid correction is possible by using - in the DCPA - a coarse grid solution operator $L_{p-1}^{-1} = (I-K_{p-1})^{-1}$ for the approximate inverse.

Combination of one relaxation step and one coarse grid correction step yield the TLA with

$$\begin{aligned} M_p^{TLA} &= (I-L_{p-1}^{-1}L_p)K_p \\ &= (I-K_{p-1})^{-1}\ (K_p-K_{p-1})K_p . \end{aligned}$$

Under suitable conditions (see HEMKER & SCHIPPERS, 1979) it can be shown that - if the repeated trapezoidal rule is used for the discretization of the integral equation - we have

$$\|M_p^{TLA}\| \le \|(I-K_{p-1})^{-1}\|\ \|(K_p-K_{p-1}\| \le C.h_p^2, \quad \text{for } p \to \infty.$$

The TLA still needs the exact solution of the discretized equation on the lower level p-1. Approximating this solution by recursive application of σ MLA iterations on lower levels we have the MLA with

$$\begin{aligned} M_p^{MLA} &= (I-(I-(M_{p-1}^{MLA})^{\sigma}L_{p-1}^{-1}L_p)K_p \\ &= M_p^{TLA} + (M_{p-1}^{MLA})^{\sigma}L_{p-1}^{-1}L_p\ K_p \\ &= M_p^{TLA} + (M_{p-1}^{MLA})^{\sigma}(K_p-M_p^{TLA}), \qquad p = 1,2,3,\ldots \end{aligned}$$

Hence,

$$\rho_p \equiv \|M_p^{MLA}\| \le \|M_p^{TLA}\| + \rho_{p-1}^{\sigma}(\|K_p\| + \|M_p^{TLA}\|.$$

From this it can be derived that, for $\sigma = 2$ and with $\rho_0 = \|M_0^{MLA}\| = \|M_0^{TLA}\|$ small enough, we have

$$\rho_p \le C\,\|M_p^{TLA}\| = \mathcal{O}(h_p^2) \text{ as } p \to \infty.$$

This is the typical behaviour of the multigrid iteration of the second kind: the finer the discretization of the analytical problem is, the faster converges the iterative process to solve the discrete system of equations.

REFERENCES

ATKINSON, K.E., *A survey of numerical methods for the solution of Fredholm integral equations of the second kind*, SIAM, 1976.

BRAKHAGE, H., *Ueber die Numerische Behandlung von Integralgleichungen nach der Quadraturformelmethode*, Num. Math., 2(1960) 183-196.

BRANDT, A., *Multi-level adaptive techniques for singular perturbation problems*, In: Numerical Analysis of Singular Perturbation Problems, P.W. Hemker & J.J.H. Miller eds., Academic Press, London, 1979.

FREDERICKSON, P.O., *Fast approximate inversion of large sparse linear systems*, Report 7-75, Lakehead University, 1975.

HACKBUSCH, W., *On the Convergence of Multigrid Iterations*, Report 79-4, Mathematisches Institut, Universität Köln, 1979.

HEMKER, P.W. & H. SCHIPPERS, *Multiple grid methods for the solution of Fredholm integral equations of the second kind*, Report NW75, Mathematisch Centrum, Amsterdam, 1979.

NICOLAIDES, R.A., *On some theoretical and practical aspects of multigrid methods*, Math. Comp. 147 (1979) 933-952.

STETTER, H.J., *The defect correction principle and discretization methods,* Num. Math. 29 (1978) 425-443.

YOUNG, D.M., *Iterative solution of oarge linear systems,* Academic Press, 1971.

(Received, November 18, 1980)