December, 1987

Rigorous Local Mode Analysis of Multigrid

Achi Brandt[1]

Department of Applied Mathematics & Computer Science
The Weizmann Institute of Science
Rehovot 76100, Israel

Abstract

Exact numerical convergence factors for the two-grid multigrid cycle can be predicted by local mode (Fourier) analysis. For general elliptic PDE systems with piecewise smooth coefficients in general domains discretized by uniform grids, it is proved that, in the limit of small meshsizes, these predicted factors are indeed obtained, provided a proper treatment is applied at and near the boundaries. That treatment, it is proved, costs negligible extra computer work and can consist of just local Kaczmarz relaxation.

Content

[1] [1] **Research supported by the Air-Force Office of Scientific Research, United States Air Force under grant numbers AFOSR-86-0126 and AFOSR-86-0127 and by the National Science Foundation under grant number NSF DMS-8704169.**

§1 Introduction

Since the earliest days of multigrid development, the "local mode analysis", based on heuristic local Fourier decomposition of the error function, has been the chief tool for the practical design, precise quantitative understanding, and even debugging, of the various multigrid processes. Although rigorously justified in very special cases only, the easily-computable predictions of that analysis have turned out precise for quite general PDE systems discretized on uniform grids, with quite general domains and boundary conditions. In several important cases, however, the predicted convergence factors were not obtained, presumably due to the influence of boundaries, which are usually not accounted for by the local analysis; domains with reentrant corners are a notorious example.

The purpose of this article is to give a general rigorous framework to the local mode analysis on one hand, and to the treatment of boundaries on the other hand. Using essentially the weakest possible assumptions, it will be proved, for general PDE systems in general domains, that the convergence factors predicted by the local mode analysis *can* be obtained. That is, the predicted factors *are* indeed obtained, provided a proper treatment, costing negligible extra work (when the meshsize is sufficiently small), is applied at and near boundaries.

The convergence factors thus proven are not just qualitative; they are *quantitatively sharp:* they are exactly obtained (or arbitrarily closely approached) by the worst local mode. By comparison, almost all other multigrid theories (see, e.g., [H] and [MMB] and references therein) give estimates which are not quantitative (containing unspecified constants) or grossly unrealistic, rendering them useless in practice (see discussion in [G, §14]). Furthermore, these other theories are either restricted to variational problems ([MMB] and references therein) or require an unknown, sufficiently large number of relaxation sweeps per cycle to guarantee convergence (thus actually analyzing much inferior algorithms). Quantitatively realistic (sometimes even sharp) two-level convergence estimates for general symmetric problems, not necessarily on grids, were derived in [AMGT], and for a very special case also in [Brs], but those estimates — unlike the present theory — cannot be improved by adding more relaxation sweeps per cycle, and cannot be generalized to V cycles.

For simplicity and clarity we first (Secs. 2–8) confine our detailed presentations to two-level one-cycle L_2 convergence factors for systems with constant

coefficients and pointwise relaxation. Extensions to variable coefficients, block relaxation, other norms (related also to other inter-grid transfer orders), many cycles (asymptotic convergence), more levels, and FMG algorithms — are then discussed in the form of comments (Secs. 9–10). The practical implications of our analysis, including powerful general ways for constructing and debugging multigrid solvers, are then reviewed (Sec. 11).

§2 Bigrid cycle

The basis and/or the model for any theoretical treatment of multigrid solvers is the analysis of the multigrid cycle in the simplest case where only two grids are involved. We assume for simplicity that these are two uniform square girds, with meshsizes h and $H = 2h$. The fine grid system of equations (no distinction is yet made between interior equations and boundary conditions) is written as

$$A^h u^h = f^h \tag{2.1}$$

where u^h and f^h are real or complex vector-valued functions on Ω^h, the intersection of the lattice $\{x = (\alpha_1, \ldots, \alpha_d)h \mid \alpha_i \text{ integers}\}$ with the bounded problem domain $\Omega \subseteq \mathbb{R}^d$. More generally, Ω^h may in fact be *staggered*; that is, $u^h = (u^{h,1}, \ldots, u^{h,q})^T$ and $f^h = (f^{h,1}, \ldots, f^{h,q})^T$, and each $u^{h,j}$ and each $f^{h,j}$ are scalar-valued functions defined on different uniform grids, $\Omega^{h,1,j}$ and $\Omega^{h,2,j}$ respectively, where each $\Omega^{h,k,j}$ is the intersection of the problem domain Ω with a translated lattice

$$\{x = (\alpha_1, \ldots, \alpha_d)h + s^{h,k,j} \mid \alpha_i \text{ are integers}\} \tag{2.2}$$

for some fixed $s^{h,k,j} \in \mathbb{R}^d$. A similar staggered grid may be introduced with meshsize $H = 2h$, and the functions defined on it will be denoted u^H and f^H. The coarse-to-fine interpolation of solutions (or approximate solutions, or correction to solutions) is denoted I_H^h; e.g., $u^h = I_H^h u^H$. The fine-to-coarse transfer of right-hand sides (or residuals) is denoted I_h^H; e.g., $f^H = I_h^H f^h$. An operator (matrix) A^H is given on grid $2h$ which approximates A^h: it may be constructed similarly to A^h by discretizing the same differential problem, or it may be defined by the Galerkin approximation $A^H = I_h^H A^h I_H^h$ (see comments in Sec.). The orders of I_H^h and I_h^H, the approximation of A^h by A^H and other assumptions will be discussed below (Sec. 6).

With this notation, the *bigrid cycle* can now be defined. Reserving the notation u^h for the exact solution of (2.1), we will replace the superscript h by other superscripts to denote various approximations to u^h. The cycle starts with a given approximation u^A, and improves it by the following three steps.

(i) *Pre-relaxation.* ν_1 relaxation sweeps are first performed using the fine grid equations (2.1). Typical is the Gauss-Seidel relaxation sweep, where the

discrete equations are scanned in some prescribed order, each one in its turn being satisfied by changing a corresponding unknown. This is natural for equations derivable from a variational (e.g., minimization) principle, where each equation indeed corresponds to one unknown. For more general cases, more general types of relaxation schemes exist. A general way for constructing good relaxation schemes for general discretized PDE systems is described in [G84, §3.7]; the description is for the *interior* relaxation, not near boundaries, but that is all one actually needs (cf. Sec. 4 below).

(ii) *Coarse grid correction* (CGC). Denoting by u^B the approximate solution obtained at the end of Step (i), and by $v^B = u^h - u^B$ the corresponding error, a coarse grid approximation to v^B, v^H, is calculated by solving the coarse grid equations

$$A^H v^H = I_h^H (f^h - A^h u^B). \tag{2.3}$$

Then v^H is used to correct the fine grid approximation:

$$u^C = u^B + I_H^h v^H \tag{2.4}$$

(iii) *Post-relaxation.* Starting with u^C, ν_2 additional relaxation sweeps are performed, yielding the final approximation u^D.

In assessing the efficiency of this cycle, one should of course disregard the work involved in solving (2.3), because in the true multigrid cycle these equations are solved approximately, by employing recursively one or two similar cycles at the coarser level.

Notation. The linear relaxation operator will be denoted by R. That is, $v^B = R^{\nu_1} v^A$ and $v^D = R^{\nu_2} v^C$, where $v^A = u^h - u^A$, $v^B = u^h - u^B$, $v^C = u^h - u^C$ and $v^D = u^h - u^D$.

The *cycle convergence factor* is defined to be

$$\lambda = \sup \frac{\|v^D\|}{\|v^A\|}, \tag{2.5}$$

where the sup is taken over all possible initial approximations $u^A \neq u^h$. The norm in (2.5) is the ℓ_2 norm, which will be used throughout most of the presentation here, although other norms may often be more appropriate (see Sec.). Our purpose is to calculate λ. Another possible purpose, to which we will refer below, is to calculate the *asymptotic convergence factor per cycle*, defined by

$$\lambda' = \limsup_{n\to\infty} (\|v^{[n]}\| / \|v^{[0]}\|)^{1/n}, \tag{2.6}$$

where $v^{[n]}$ is the error after n applications of the cycle, so in particular $v^{[0]} = v^A$ and $v^{[1]} = v^D$.

§3 Mode analysis in entire space

In case $\Omega = \mathbb{R}^n$ (and hence also in case of a rectangular domain with periodic boundary conditions, where the problem can be extended to the entire space $\mathbb{R}^n$), and assuming "constant coefficient" operators A^h and A^H and a "consistently ordered" relaxation scheme (terms to be defined below), the convergence factors λ and λ' can easily be calculated by a mode (Fourier) analysis. That is, the initial error v^A, and similarly the error v^h at any other stage, can be written as a linear combination (integral; or sum, in case of periodic boundary conditions) of Fourier modes

$$v^h_\alpha = \int_{|\theta| \le \pi} \hat{v}^h(\theta) e^{i\theta\alpha}\, d\theta, \tag{3.1}$$

and the change of the coefficients $\hat{v}^h(\theta)$ under each of the processes in the cycle can explicitly be calculated, yielding an explicit calculation of λ (or λ').

The notation in (3.1) is as follows: $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a vector of integers, $v^h_\alpha = (v^{h,1}_\alpha, \ldots v^{h,q}_\alpha)^T$ with $v^{h,j}_\alpha = v^{h,j}(\alpha h + s^{h,1,j})$, $\theta = (\theta_1, \ldots, \theta_d)$, $\theta\alpha = \theta_1\alpha_1 + \cdots + \theta_d\alpha_d$, $|\theta| = \max(|\theta_1|, \ldots, |\theta_d|)$ and

$$\hat{v}^h(\theta) = (2\pi)^{-d} \sum_\beta v^h_\beta e^{-i\theta\beta}, \tag{3.2}$$

$\sum_\beta$ denoting summation over the entire grid of integer vectors $\beta = (\beta_1, \ldots, \beta_d)$. As implied by (3.2), $\hat{v}^h(\theta) = \hat{v}^h(\theta_1, \ldots, \theta_d)$ can naturally be extended as a 2π-*periodic* function, i.e.,

$$\hat{v}^h(\theta) = \hat{v}^h(\theta_1 + 2\pi, \theta_2, \ldots, \theta_d) = \cdots = \hat{v}^h(\theta_1, \ldots, \theta_{d-1}, \theta_d + 2\pi) \tag{3.3}$$

for any $\theta \in \mathbb{R}^d$. To be precise, the integration meant in (3.1), and similarly below, is over one cell of this period, e.g. the cell

$$-\pi \le \theta_j < \pi, \qquad (j = 1, \ldots, d). \tag{3.4}$$

We assume that the decomposition (3.1) exists, as well as the Parseval identity

$$\int_{|\theta| \le \pi} |v^h(\theta)|^2 d\theta = (2\pi)^{-d} \sum_\alpha |v^h_\alpha|^2.$$

This is true for a wide class of error functions. We do not prove it here, because in this section the development is purely formal; a proof will be given later (Sec. 8.1), when dealing with real, bounded domains Ω.

We use the notation A^h and A^H for the discrete fine-grid and coarse-grid operators, respectively, including both interior and boundary operations. In case of the infinite domain they of course coincide with the *interior operators*, which

we denote by L^h and L^H, respectively. That L^h has "constant coefficients" means that it has the (block-Toeplitz) form

$$(L^h u^h)_\alpha = \sum_\gamma a^h_\gamma(h) u^h_{\alpha+\gamma}, \tag{3.5}$$

where $\sum_\gamma$ is a summation over integer vectors $\gamma = (\gamma_1, \ldots, \gamma_d)$, and each $a^h_\gamma(h)$ is a $q \times q$ matrix which explicitly depends on the meshsize h (using it in the denominator of divided differences). Hence, with the Fourier decomposition (3.1) for v^h, it is easy to see that if $r^h = L^h v^h$ and

$$r^h_\alpha = \int_{|\theta| \le \pi} \hat{r}^h(\theta) e^{i\theta\alpha}\, d\theta \tag{3.6}$$

then $\hat{r}^h(\theta) = \hat{L}^h(\theta, h)\hat{v}^h(\theta)$, where

$$\hat{L}^h(\theta, h) = \sum_\gamma a^h_\gamma(h) e^{i\theta\gamma}. \tag{3.7}$$

$\hat{L}^h(\theta, h)$, which is an easily computed (easily programmed) $q \times q$ matrix of functions of θ, is called the *symbol* of L^h.

The symbol $\hat{L}^H(\theta, H)$ of the coarse grid operator can be similarly defined (applied to *coarse* grid functions v^H, and hence to Fourier modes $\exp(i\theta x/H)$, instead of the modes $\exp(i\theta/h)$ used in (3.1)). Often, L^H and L^h are identical discretizations of the same differential operator, using only different meshsizes, in which case $\hat{L}^H(\theta, H) = \hat{L}^h(\theta, 2h)$. Note that when L^H is applied to a *fine* grid function v^h (with the *fine* grid expansion (3.1)), its symbol is $\hat{L}^H(2\theta, H)$.

We thus see that the operator $A^h = L^h$ does not couple different modes: For each θ, the Fourier coefficient $\widehat{L^h v^h}(\theta)$ depends only on $\hat{v}^h(\theta)$ for the same θ. This property also holds for the operation of relaxation, provided the relaxation scheme is "consistently ordered", i.e., for any α and γ the point α is relaxed *after* the point $\alpha - \gamma$ if and only if it is relaxed *before* $\alpha + \gamma$. We will however soon extend the class of relaxation schemes treated by our analysis to include some important schemes which are not strictly consistently ordered, such as red-black schemes.

On grid $H = 2h$, the Fourier mode $\exp(i\theta\alpha)$ "aliases" (coincides) with any other mode $\exp(i\theta'\alpha)$ for which $\theta = \theta' \pmod{\pi}$, i.e., for which each $(\theta_j - \theta'_j)/\pi$ is an integer, $(j = 1, \ldots, d)$. Such modes, or such θ and θ', we will call *harmonics* of each other. Every intergrid transfer, either I^H_h or I^h_H, must couple each component with all its harmonics. If the transfer has constant coefficients (i.e., it repeats itself at each coarse grid cell), as we will assume, then it will couple *only* harmonics. We will therefore consider simultaneously each set of harmonics in the range (3.4).

For this purpose, let $\tau^j_1 \ldots \tau^j_d$ be the binary representation of the integer j, $0 \le j < 2^d$, i.e., $j = \sum_{k=1}^d 2^{d-k}\tau^j_k$, and let $\tau^j = (\tau^j_1, \ldots, \tau^j_d)\pi$. Each set of

harmonics in the range (3.4) includes one frequency θ in the range

$$-\frac{\pi}{2} \le \theta_k < \frac{\pi}{2}, \qquad (k = 1, \ldots, d) \tag{3.8}$$

which we will call *the lowest harmonic*. For each such θ, its set of harmonics in the range (3.4) is the set of $D = 2^d$ components

$$\{\theta^j = \theta + \tau^j \pmod{2\pi}; \quad j = 0, 1, \ldots, D-1\}, \tag{3.9}$$

including the lowest harmonic $\theta = \theta^0$ itself.

On the coarse grid each θ^j appears as θ^0. Hence, operating with a constant-coefficient fine-to-coarse transfer I_h^H on any fine-grid residual such as r^h in (3.6) gives

$$(I_h^H r^h)_\alpha = \int^0 \sum_{j=0}^{D-1} \hat{I}_h^H(\theta^j) \hat{r}^h(\theta^j) e^{i\theta^0 \alpha} \, d\theta^0, \tag{3.10}$$

where $\int^0$ denotes, here and below, integration over the domain (3.8), with θ^j relating to θ^0 through (3.9), and $\hat{I}_h^H(\theta)$ is a $q \times q$ matrix of easily programmed functions of θ, called the *symbol* of I_h^H. Similarly, the correction interpolation I_H^h will transfer the coarse grid solution

$$v_\alpha^H = \int^0 \hat{v}^H(\theta^0) e^{i\theta^0 \alpha} \, d\theta^0 \tag{3.11}$$

to the fine grid correction

$$(I_H^h v^H)_\alpha = \int^0 \sum_{j=0}^{D-1} \hat{I}_H^h(\theta^j) \hat{v}^H(\theta^0) e^{i\theta^j \alpha} \, d\theta^0, \tag{3.12}$$

where the $q \times q$ matrix of functions $\hat{I}_H^h(\theta)$ is called the symbol of I_H^h. Note that if I_h^H is the adjoint of I_H^h, as is often the case, then $\hat{I}_h^H(\theta) = \hat{I}_H^h(\theta)^\dagger$ (superscript $\dagger$ denoting conjugate transposition). Most often both these symbols are diagonal. For I-order multipolynomial interpolation, for example,

$$\hat{I}_H^h(\theta)_{k,\ell} = \delta_{k,\ell} \prod_{i=1}^{d} \varphi_I(\cos\theta_i), \tag{3.13}$$

where $\varphi_2(\xi) = (1+\xi)/2$, $\varphi_4(\xi) = (2 + 3\xi - \xi^3)/4$, etc.

Since harmonics are coupled anyway, we will allow them to be coupled also by the relaxation process. We thus extend the concept of consistently ordered

relaxation to any relaxation operator R such that, operating with it on the error function (3.1) will give an error function of the form

$$(Rv^h)_\alpha = \int^0 \sum_{j,k=0}^{D-1} \hat{R}_j(\theta^k)\hat{v}^h(\theta^k)e^{i\theta^j\alpha}\,d\theta^0. \tag{3.14}$$

This will include red-black (and even 2^d-colored) relaxation schemes. $\hat{R}_j(\theta)$ are easily computed $q \times q$ matrices (see [MOC, §3.1] or [G84, §1.2] for a simple example).

We can now describe, in terms of Fourier transforms in which harmonics are being blocked together, the entire bigrid cycle. To this end we introduce the following *block-matrix notation:* for any fine-grid error function v^h (and similarly for any residual function; each being a *vector* of q functions), with the Fourier transform (3.1), $\check{v}^h(\theta^0)$ will denote the vector (of length qD) $(\hat{v}^h(\theta^0)^\dagger, \ldots, \hat{v}^h(\theta^{D-1})^\dagger)^\dagger$. In particular, the initial error v^A has the block-Fourier decomposition

$$v^A_\alpha = \int^0 E_\alpha \check{v}^A(\theta)e^{i\theta\alpha}\,d\theta, \tag{3.15}$$

where E_α is the $q \times qD$ matrix

$$E_\alpha = (e^{i\tau^0\alpha}I_q, \ldots, e^{i\tau^{D-1}\alpha}I_q), \tag{3.16}$$

I_q being the $q \times q$ identity matrix. The error at the cycle end has a similar decomposition

$$v^D_\alpha = \int^0 E_\alpha \check{v}^D(\theta)e^{i\theta\alpha}\,d\theta, \tag{3.17}$$

and by the definition of the cycle we obtain the relation

$$\check{v}^D(\theta) = M(\theta)\check{v}^A(\theta), \qquad (-\frac{\pi}{2} \le \theta_i < \frac{\pi}{2}, \quad i = 1, \ldots, d) \tag{3.18}$$

where $M(\theta)$ is the $qD \times qD$ matrix

$$M(\theta) = \check{R}(\theta)^{\nu_2}[\check{I} - \check{I}^h_H(\theta)\check{L}^H(\theta)^{-1}\check{I}^H_h(\theta)\check{L}^h(\theta)]\check{R}(\theta)^{\nu_1}. \tag{3.19}$$

The block matrices in (3.19) are defined by

	full matrix dimension
$\check{R}(\theta)_{JK} = \hat{R}_{J-1}(\theta^{K-1})$	$qD \times qD$
$\check{L}^h(\theta)_{JK} = \delta_{JK}\hat{L}^h(\theta^{J-1}, h)$	$qD \times qD$
$\check{I}^H_h(\theta)_J = \hat{I}^H_h(\theta^{J-1})$	$q \times qD$
$\check{L}^H(\theta) = \hat{L}^H(2\theta, H)$	$q \times q$
$\check{I}^h_H(\theta)_K = \hat{I}^h_H(\theta^{K-1})$	$qD \times q$
$\check{I}_{JK} = \delta_{JK}I_q$	$qD \times qD$

where $\theta^0, \ldots, \theta^{D-1}$ are related to θ by (3.9), and where J and K are *block* indices, pointing to the $q \times q$ block at the J-th column and K-th row ($J, K = 1, \ldots, D$ whenever appearing; the full matrix dimension shown on the right also indicates their range).

From (2.5), (2.6), (3.15), (3.17), (3.18) and the Parseval identity it formally follows that

$$\lambda = \sup_{\theta \neq 0} \|M(\theta)\| \tag{3.20}$$

and

$$\lambda' = \sup_{\theta \neq 0} \sigma(M(\theta)), \tag{3.21}$$

where $\sigma(M)$ is the spectral radius of M (i.e., its largest absolute eigenvalue), $\|M\|$ is its ℓ_2 operator norm (i.e., $\|M\| = \sigma(MM^T)^{1/2}$) and each sup is over the range (3.8), with $\theta = 0$ being excluded. This exclusion is important since most often $\check{L}^H(0)$ is singular; the sup is normally still finite since $\check{I}_h^H(0)\check{L}^h(0)$ is suitably singular (rank deficient) too. For some bigrid cycles, λ may be infinite; on a bounded domain, as we will see, such cycles can still be used, provided that λ' is still finite.

It is easy to program the matrix function $M(\theta)$, hence to calculate (3.20) and (3.21). Our task is to prove that, for sufficiently small meshsizes and with a proper treatment of boundaries, the values of λ and λ' (defined by (2.5) and (2.6)) on any *bounded* domain are still given (or approximated as closely as one wishes) by (3.20) and (3.21).

§4 Treatment of boundaries: general approach

It is well known that the efficiency of the bigrid (and other multigrid) cycles may strongly be affected by the shape of the boundary curve (e.g., existence of reentrant corners), by the type and coefficients of the boundary conditions, by the boundary position relative to both the fine grid and the coarse grid, and by the discretization and solution processes (relaxation and inter-grid transfers) employed at and near the boundary. With this enormous variety, we believe it is unproductive in a general theory to rigorously analyze any one or other particular boundary situation. We will instead show that, in the limit $h \to 0$, the details of the boundary processes are never important, since, on one hand, they employ negligible amounts of computations, and, on the other hand, they can in a simple way always be *chosen* so that the overall efficiency (e.g., the convergence factor per cycle) is just the efficiency dictated by the interior processes.

Aside from simplicity and generality, two related reasons lead to this approach. First, we aim at an *exactly quantitative* analysis, i.e., calculating the actual numerical value (not just upper bounds) of λ and λ'. This would be too difficult to do when all the details of complicated boundary situations should be

taken into account. Secondly, from a practical point of view, our approach gives the more important information: it tells us what efficiency one *should* be getting once the boundary processes have been *properly* adjusted. The analysis below also shows *one general way* of making this adjustment, although in practice, in each particular case some other ways may be more convenient or more effective at large values of h. (See the treatment of reentrant corners in [Bai, §4] and in [ZN]. For generality, in our treatment here, we do *not* exploit any smoothness properties of the boundary shape and the boundary operators.)

Thus, our approach is to allow the analyzed bigrid cycle to be modified near boundaries (and similarly also near other singular curves, such as interfaces, material discontinuities, etc.; cf. Sec.) provided the work involved is negligible. The general way we propose to modify the cycle is to add a certain number of Kaczmarz relaxation passes over the boundary conditions and over the interior equations in some small neighborhood of the boundary (see details in Sec. 7). We will therefore present now Kaczmarz relaxation and its relevant properties.

§5 Kaczmarz relaxation. Partial sweeps

Consider the general system of (real or complex) equations

$$Au = f \quad \text{or} \quad \sum_{j=1}^{n} a_{ij} u_j = f_i, \qquad (i = 1, \ldots, n). \tag{5.1}$$

Given any approximation u^*, a *Kaczmarz relaxation step for the i-th equation* is defined as the replacement of u^* by the vector closest to it on the i-th hyperplane (the hyperplane of solutions to the i-th equation). This means that each u_k^* is replaced by $u_k^* + \beta_i \overline{a}_{ik}$, where $\beta_i = r_i^* / \sum_{j=1}^{n} |a_{ij}|^2$ and $r_i^* = f_i - \sum_{j=1}^{n} a_{ij} u_j^*$. This r_i^*, the residual of the i-th equation just before relaxing it, is called the i-th *dynamic* residual. Note that the residual of the equation just *after* relaxing it is zero, of course.

Suppose a Kaczmarz relaxation pass is made over the first m equations ($1 \le m \le n$); i.e., for each of those equations in its turn in the natural order, a Kaczmarz relaxation step is performed. Denote the solution vectors before and after this relaxation pass by u^0 and u^1, respectively; the corresponding error vectors by $v^0 = u - u^0$ and $v^1 = u - u^1$; the corresponding residual vectors by $r^0 = f - Au^0 = Av^0$ and $r^1 = f - Au^1 = Av^1$; and the corresponding *normalized* residual vectors by $\tilde{r}^0$ and $\tilde{r}^1$, where

$$\tilde{r}_i^s = r_i^s \Big/ \left(\sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}, \qquad (i = 1, \ldots, n; \quad s = 0, 1, *). \tag{5.2}$$

Also, denote by δ_i^* the ℓ_2 norm of the solution change while relaxing the i-th equation, and note that

$$\begin{aligned} |\delta_i^*|^2 &= |\beta_i|^2 \sum_{k=1}^{n} |a_{ik}|^2 \\ &= |\tilde{r}_i^*|^2, \qquad (i = 1, \ldots, m), \end{aligned} \tag{5.3}$$

where $\tilde{r}_i^*$ is the *normalized dynamic* residual (defined in (5.2)). Using this notation and the ℓ_2 norm $\|v\|^2 = \sum_i |v_i|^2$, we can formulate the following general property of Kaczmarz relaxation.

Theorem 5.1

$$\begin{aligned} \|v^0\|^2 - \|v^1\|^2 &= \sum_{i=1}^{m} |\delta_i^*|^2 = \sum_{i=1}^{m} |\tilde{r}_i^*|^2 \\ &\geq \max\left(\gamma_0 \sum_{i=1}^{m} |\tilde{r}_i^0|^2, \quad \gamma_1 \sum_{i=1}^{m} |\tilde{r}_i^1|^2 \right) \end{aligned} \tag{5.4}$$

where

$$\begin{aligned} \gamma_0 &= [(1+\gamma_-)(1+\gamma_+)]^{-1} \\ \gamma_1 &= (\gamma_-\gamma_+)^{-1} \\ \gamma_- &= \max_{1\leq i\leq m} \left(\sum_{j=1}^{i-1} \sum_{\lambda=1}^{n} |a_{i\lambda}\overline{a}_{j\lambda}| \right) \Big/ \sum_{\lambda=1}^{n} |a_{i\lambda}|^2 \\ \gamma_+ &= \max_{1\leq i\leq m} \left(\sum_{j=i+1}^{m} \sum_{\lambda=1}^{n} |a_{i\lambda}\overline{a}_{j\lambda}| \right) \Big/ \sum_{\lambda=1}^{n} |a_{i\lambda}|^2 \end{aligned} \tag{5.5}$$

Proof. The proof is analogous to the proof of Theorem 3.3 in [AMGT], which is the special case $m = n$. To see that the proof carries over to the case $m < n$, observe that the proofs of Theorems 3.1 and 3.2 in [AMGT] are easily modified to this case. ∎

The theorem essentially says that the convergence is fast as long as the average normalized residual of the relaxed equations is comparable to the average error, averages being meant in the ℓ_2 sense. The coefficients γ_0 and γ_1 can be interpreted as rough measures for the independence of the relaxed equations.

Remark 5.1. Note that Kaczmarz relaxation, as well as v^0, v^1, $\tilde{r}^*$, $\tilde{r}^0$ and $\tilde{r}^1$ appearing in (5.4), remain unchanged when each equation is rescaled (multiplied through by a constant). We can thus decrease $\gamma_\pm$ (thereby increasing γ_1 and γ_2) in the theorem to any values obtainable by such rescaling. It is then easy to see that

for any local operator all γ_i are $O(1)$. For example, this is obtained by rescaling so that $\sum_{j=1}^{n} |a_{ij}|^2 = 1$ for all i. In addition, γ_i will be finite for many *non*-local operators.

Remark 5.2. Even more important is the correct scaling of the *unknowns*. This scaling actually changes the Kaczmarz steps. A typical example is the Stokes system of equations ($\Delta U = P_x$, $\Delta V = P_y$ and $U_x + V_y = 0$, where Δ is the Laplacian and subscripts denote partial differentiations. To fit the framework of this paper this system should be discretized h-elliptically. This can be done either on a staggered grid, as in [G84, §18.2], or on a non-staggered grid, as in [CM3, App. C]). Each variable P_α in this system should conceptually be replaced by hP_α before Kaczmarz relaxation is applied, basically because only in terms of this rescaling the CGA assumption can hold (see Remark 6.1 below).

§6 Assumptions

Listed below are all our assumptions about the mathematical properties of the various multigrid ingredients: relaxation, A^h, A^H, I_h^H and I_H^h. The discussion will explain that each of these assumptions is more or less *necessary*; the main content of our theory is of course to show, in subsequent sections, that the assumptions are *sufficient* — sufficient for achieving the predicted convergence factor (3.20). Except perhaps for the CGA (see discussion in Sec. 6.1), all other suppositions will be easily verifiable in any case of interest. They are all *qualitative* assumptions, in the sense that their main constants are arbitrary and unspecified. Indeed, the very point of this article is that such qualitative and necessary assumptions yield a *quantitative* and even *precise* prediction of convergence factors.

An *un*necessary assumption which for simplicity we do introduce at this first part of our presentation is that all interior processes (at distance greater than $O(h)$ from the boundary) have **constant coefficients**. This includes L^h, L^H, I_h^H and I_H^h, as well as the **consistent ordering** of relaxation (extended as in Sec. 3). It is indeed only under this assumption that the local mode analysis is straightforwardly *defined*. The removing of this assumption, and the corresponding extension of the local mode analysis, are discussed in Sec. .

General notation. The letter C will be used to denote any constant, not necessarily the same on different occurrences, independent of h. The notation C_ℓ will be used in case the constant depends on the integer ℓ.

The more important (less obvious) assumptions are described first (Secs. 6.1 and 6.2).

6.1 Coarse grid approximation (CGA)

A necessary condition for the multigrid process to work properly is that errors which would converge slowly under *any* relaxation scheme must be well approximated on the coarse grid. The slower an error is reduced by any possible relaxation scheme, the better must its coarse grid approximation be. When such a condition is *not* satisfied, convergence factor of the bigrid cycle cannot be made uniformly (for all meshsizes) as small as one wishes by adding more relaxation sweeps (increasing $\nu_1 + \nu_2$). Moreover, if such a condition is unsatisfied at all multigrid levels, the V cycle cannot be expected to produce convergence factors bounded away from 1 independently of the number of levels.

The basic relaxation schemes are the *point-by-point* ones — in the sense that *block* schemes, such as line relaxation, employ at each of their steps a solution process which could itself be multigridded, using a point-by-point scheme for its relaxation, hence the entire process may be interpreted as based on a point relaxation scheme with semi coarsening (see [G, §4.2.1]). For simplicity *we will therefore refer below to point schemes only* and defer discussion of block schemes to later comments (see Sec.). Now, any point-by-point relaxation scheme introduces changes to the solution based on the size of the local residuals relative to the size of the coefficients of the corresponding equations, hence *any* such scheme *must* exhibit slow convergence when an error vector v develops for which the normalized residuals $\tilde{r}$ are small, i.e., when $\|\tilde{r}\| \ll \|v\|$.

Unlike the *geometric notation* of Secs. 2 and 3, we have used in Secs. 4 and 5 the *algebraic notation*, where the unknowns u are arranged in one long vector, and the fine-grid operator A^h is correspondingly arranged as a big matrix A, with $(Au)_i = \Sigma a_{ij} u_j$. Keeping this notation, and motivated by the above discussion, we now introduce the *normalized operator* $\tilde{A}^h$, corresponding to the matrix $\tilde{A}$ defined by

$$(\tilde{A}u)_i = (Au)_i / (\sum_j |a_{ij}|^2)^{1/2}. \tag{6.1}$$

Thus, $\tilde{A}^h v = \tilde{r}$, and the above condition, which should be satisfied by any multigrid algorithm employing point relaxation, can be put as follows.

Coarse Grid Approximation (CGA) condition. *Denoting by v^B and v^C the algebraic error vectors before and after the coarse grid correction, respectively, for any $\epsilon > 0$ there exists $\delta = \delta(\epsilon)$, independent of the meshsize, such that, if $\|\tilde{A}^h v^B\| \le \delta \|v^B\|$ then $\|v^C\| \le \epsilon \|v^B\|$.*

We will *assume* this condition to hold. This will enable us to avoid dealing with many different cases and with details of boundary conditions. Indeed, the main point in this paper is to show that *quantitatively sharp* convergence factors (λ, λ') can be derived from *qualitative* assumptions, such as this CGA assumption. The CGA, and in fact much stronger conditions, are normally assumed or proved

in any multigrid theory, without having formerly yielded any realistic convergence constants at all, let alone the sharp constants of the local mode analysis. CGA proofs can usually be derived for any discretization scheme for which a *convergence* proof exists, containing some estimates of the fine-grid truncation errors in terms of the continuum-solution smoothness: by analogy, the coarse-grid truncation errors can be estimated in terms of the fine-grid error smoothness. It is in those proofs that one needs to deal with details of boundary and near-boundary situations (cf. Sec. 4).

Note that the CGA condition conveniently unifies boundary and interior equations. By using the *normalized* operator we put them all on the same footing. Also note that we have required a very weak form of the condition, allowing the decrease in δ (as a function of ϵ) to be arbitrarily fast. In fact, the preceding arguments indicate that this is the weakest possible condition when pointwise relaxation is used, in the sense that without it one cannot indefinitely increase the convergence factor of the bigrid cycle (uniformly for all h) by increasing ν, nor can one generally expect h-independent convergence of the V multigrid cycle.

Remark 6.1. In case of *systems* ($q > 1$, in the notation of Sec. 2), the CGA can often hold only when the different functions are properly scaled relative to each other. In case of the Stokes system (see Remark 5.2), for example, in terms of the original set of functions (U, V and P) the CGA condition cannot be satisfied: For instance, an error v^B which consists of no error in U and V and a highly oscillating error in P cannot be much reduced by any coarse grid correction (because it is highly oscillatory), although it does satisfy $\|\tilde{A}^h v^B\| \leq O(h)\|v^B\|$. The CGA will only be satisfied when the function P is replaced by hP. In some other systems, more involved changes of variables, and perhaps also recombinations of equations, may be needed before the CGA can hold.

6.2 Inter-grid transfer orders

Another necessary condition for a multigrid cycle to work properly is that suitable orders are used for the coarse-to-fine correction interpolation operator I_H^h and for the fine-to-coarse residual transfer I_h^H. By "work properly" we mean that the convergence is not degraded by CGC amplifications of high frequencies. As explained in Sec. 4.3 of [G], this implies a set of necessary rules. Our rigorous local mode analysis will in fact prove those rules to be also sufficient.

The rules deal with *interior* (not boundary) interpolations and depend on the *interior* finite-difference operators L^h and L^H. Assuming that (2.1) in the interior approximates a system of q differential equations in q unknown functions, we denote by m_{ij} the highest order of differentiation of the j-th unknown function in the i-th equation ($i, j = 1, \ldots, q$). This means, in terms of the notation in Sec. 3, that

$$\hat{L}^h(\theta)_{i,j} \leq Ch^{-m_{ij}}. \tag{6.2}$$

The order of the i-th equation is $m_{i*} = \max_{1\le j\le q} m_{ij}$.

We assume for simplicity that the residuals of each of the q equations are transferred to the coarse grid separately from those of other equations, and we denote the order of transfer of the i-th equation by m_i $(i = 1, \ldots, q)$. By this we mean that

$$\hat{I}_h^H(\theta^0 + \tau^k)_{ij} = \delta_{ij} O(|\theta^0|^{m_i}), \qquad (k = 1, \ldots, D-1). \tag{6.3}$$

Note that $k = 0$, for which $\tau^k = 0$, is excluded from (6.3).

In our present simple demonstration we are interested in λ, the L_2 convergence factor of one bigrid cycle. In this case the inter-grid transfer rules are reduced to the simple requirement

$$m_i > m_{i*}. \tag{6.4}$$

Indeed, it is easy to see that $O(1)$ high-frequency errors, in the j-th function before the CGC, contribute to the right-hand side of the i-th coarse-grid equation (and hence also to the error at the end of the cycle) smooth components whose magnitude is $O(h^{m_i - m_{ij}})$. Therefore, to ensure that the cycle reduction factor cannot be constrained by this type of error, we must assume (6.4). We could relax (6.4) by using error norms other than L_2 (see Sec. 10.1) or by considering the *asymptotic* convergence factor λ' instead of λ (see Sec.), but for simplicity of the first presentation we assume for now (6.4).

6.3 Properties of A^h

It will be assumed that A^h is a **local** operator. Namely, if we relate our algebraic notation to geometric locations through

$$(A^h u^h)(x_i) = \sum_j a_{ij} u^h(y_j) \tag{6.5}$$

then

$$a_{ij} = 0 \qquad \text{for} \qquad |x_i - y_j| > Ch. \tag{6.6}$$

For the interior operator L^h, defined by (3.5), this means that $a_\gamma^h(h) = 0$ for $|\gamma| > C$.

More precisely, it is enough to assume a somewhat weaker assumption. It will only be needed that γ_1 is finite (see Remark 5.1) and that (6.7) in the following lemma is satisfied.

Lemma 6.1. *If A^h is a local operator, then for any vector-valued grid function v and scalar continuum function φ*

$$\|\tilde{A}^h(\varphi v) - \varphi \tilde{A}^h v\| \le C\|v\| \max_{|x-y|\le Ch} |\varphi(x) - \varphi(y)|. \tag{6.7}$$

where $\tilde{A}^h$ is the normalized operator (cf. (6.1)).

Proof. By (6.5) and then (6.6) and the Cauchy-Schwarz inequality,

$$\left|[A^h(\varphi v)](x_i) - \varphi(x_i)A^h v(x_i)\right|^2 = \left|\sum_j a_{ij}[\varphi(y_j) - \varphi(x_i)]v(y_j)\right|^2$$
$$\leq \max_{|x-y|\leq Ch} |\varphi(x) - \varphi(y)|^2 \sum_j |a_{ij}|^2 \sum_{|x_i - y_j|\leq Ch} |v(y_j)|^2.$$

Hence, dividing through by $\sum_j |a_{ij}|^2$ and then summing over i, (6.7) is obtained, since each $|v(x_j)|^2$ will appear in the sum in at most $q(2C+1)^d$ terms. ∎

It will also be assumed that the interior operator L^h **approximates a differential operator** (of the first or higher order). More precisely, defining the normalized interior operator $\tilde{L}^h$ as the interior part of the normalized operator $\tilde{A}^h$ (cf. (6.1) and (3.5)), and denoting its symbol by $\hat{\tilde{L}}^h(\theta)$ (cf. (3.7)), it will be assumed that

$$\|\hat{\tilde{L}}^h(\theta)\| \to 0 \qquad \text{as} \qquad \max(h, |\theta|) \to 0, \tag{6.8}$$

where, as usual, $\|\cdot\|$ is the ℓ_2 matrix norm. Indeed, it is easy to see that if L^h is a difference approximation to a $q \times q$ differential operator (where each individual discrete equation approximates one of the q differential equations), then

$$|\hat{\tilde{L}}^h(\theta)_{ij}| \leq C(|\theta|^{m_{ij}} h^{m_{i*} - m_{ij}} + h^{m_{i*}}). \tag{6.9}$$

For some (e.g., Hermitian) finite element formulations, some of the unknown grid functions which do not correspond to continuum unknowns may have to be properly scaled for (6.8) to hold. Anyway, we use (6.8) only to show that the contribution of v^f to $\overline{v}^d$ is small (see Sec. 8.4), which should be true for any reasonable type of approximation.

6.4 Stability

An obvious requirement that should be imposed is the stability of all the cycle processes in the norms of interest — L_2 norms in the present theory. Thus, we will assume that, for any fine-grid error v^h and residual r^h and for any coarse-grid v^H and r^H, the following holds.

$$\|Rv^h\| \leq C\|v^h\| \tag{6.10}$$
$$\|(A^H)^{-1} I_h^H A^h v^h\| \leq C\|v^h\| \tag{6.11}$$
$$\|I_h^H r^h\| \leq C\|r^h\| \tag{6.12}$$
$$\|(A^H)^{-1} r^H\| \leq C\|r^H\| \tag{6.13}$$
$$\|I_H^h v^H\| \leq C\|v^H\| \tag{6.14}$$

In addition we will also assume stability of relaxation in terms of the *residual* L_2 norm, i.e.,

$$\|A^h R v^h\| \le C\|A^h v^h\|. \tag{6.15}$$

For explicit processes, such as A^h, I_h^H, I_H^h and Jacobi-type relaxation, these stability requirements are easily established. For non-explicit relaxation schemes, such as Gauss-Seidel and Kaczmarz, stability (6.10) or (6.15) can usually easily be checked by a version of local mode analysis (applied to the interior process, and adopting if necessary an approach similar to the one described in Sec. 4 above for handling boundaries): First stability is checked for the marching, point after point, within one line; then for the marching, line after line, within a plane (relating *line* Fourier decompositions of the errors); etc. It is, incidentally, quite important to check this stability, because sometimes in the search for schemes with minimal smoothing factors one can easily run into unstable ones.

Usually then, the only stability requirements which are not easily verified are (6.11) and (6.13). A vast literature treats the latter, so here we can indeed simply *assume* it. Condition (6.11) is listed here only for convenience: it is actually implied by the CGA assumption. (Because, if v^h violates (6.11) with sufficiently large C, a suitably small multiple of it can be added to v^B in the CGA condition and cause the latter to be violated.)

The stability assumptions together with the CGA assumption imply the following.

Lemma 6.2. *For any $\epsilon > 0$ and $c_* > 0$ there exists $\delta_1(\epsilon, c_*) > 0$ such that, if the error at the beginning of the cycle satisfies*

$$\|v^A\| \le c_* \tag{6.16}$$

and

$$\|\tilde{A}^h v^A\| \le \delta_1(\epsilon, c_*), \tag{6.17}$$

then the error at the end of the cycle satisfies

$$\|v^D\| \le \epsilon. \tag{6.18}$$

Proof. The stability requirements (6.10)–(6.15) imply that there exist h-independent constants C_1, C_2, C_3 and C_4 such that $\|v^B\| \le C_1\|v^A\|$, $\|A^h v^B\| \le C_2\|A^h v^A\|$, $\|v^D\| \le C_3\|v^B\|$ and $\|v^D\| \le C_4\|v^C\|$. Define

$$\delta_1(\epsilon, c_*) = \frac{\epsilon}{C_2 C_3}\delta(\epsilon_1), \quad \text{where} \quad \epsilon_1 = \frac{\epsilon}{c_* C_1 C_4}$$

and where $\delta(\epsilon_1)$ is the function defined by the CGA condition (cf. Sec. 6.1). If $\|v^B\| \le \epsilon/C_3$ then (6.18) trivially follows. If not, then by (6.17)

$$\|\tilde{A}^h v^B\| \le C_2\|\tilde{A}^h v^A\| \le C_2\delta_1 \le \frac{C_2 C_3 \delta_1}{\epsilon}\|v^B\| = \delta(\epsilon_1)\|v^B\|.$$

By the CGA assumption it thus follows that $\|v^C\| \le \epsilon_1 \|v^B\|$, and hence

$$\|v^D\| \le \epsilon_1 C_4 \|v^B\| \le \epsilon_1 C_1 C_4 \|v^A\| \le \epsilon_1 C_1 C_4 c_* = \epsilon.$$

∎

6.5 Mode-analysis expressions

For each explicit and local process, each term in its associated symbol (cf. Sec. 3) is of course a linear combination of trigonometric functions of the form

$$e^{i\theta\alpha} = e^{i(\alpha_1\theta_1 + \cdots + \alpha_d\theta_d)}, \qquad |\alpha| = \max_j |\alpha_j| \le C. \tag{6.19}$$

This is therefore true for each term in the matrices $\check{L}^h(\theta)$, $\check{I}_h^H(\theta)$, $\check{L}^H(\theta)$ and $\check{I}_H^h(\theta)$, and also $\check{R}(\theta)$ in case R is a Jacobi-type relaxation. For more general relaxation schemes, $\check{R}(\theta)$ can be expressed as a product of such matrices (whose terms are linear combinations of trigonometric functions) and inverses of such matrices. Hence each term in $\check{R}(\theta)$ is a rational function of trigonometric functions (6.19). We will assume each of those terms to be *bounded*. This assumption, which can easily be checked and trivially holds in all familiar schemes, is equivalent to requiring

$$\|\check{R}(\theta)\| \le C, \qquad (\text{uniformly in } \ |\theta| \le \frac{\pi}{2}), \tag{6.20}$$

which largely overlaps assumption (6.10) above. (More precisely: (6.20) implies (6.10) in the interior. The latter, however, unlike (6.20), would allow $\|\check{R}(\theta)\|$ to be unbounded for $|\theta| \to 0$ in a way which is not characteristic to usual relaxation schemes. The theory below allows a condition weaker, but more complicated to state, than (6.20).)

A more delicate condition should be required from $\check{L}^H(\theta)$. Since its inverse will be used, we should roughly require that $\check{L}^H(\theta)$ is nonsingular for $\theta \ne 0$, which is a sort of **ellipticity** requirement. More precisely, we can use the weaker condition

$$\det \check{L}^H(\theta) \ne 0 \qquad \text{for} \qquad |\theta| \ge h^{1-\sigma_*} \tag{6.21}$$

with a certain sufficiently small $\sigma_* > 0$. This condition expresses *ellipticity on scale* H^{σ_*} (see [G, §2]). Such conditions are easily checked and trivially satisfied by all discretized elliptic systems; but we can substantially further weaken them here.

Observe that $\det \check{L}^H(\theta)$ is a polynomial in trigonometric functions (6.19), hence, for small $|\theta|$ it is approximately a polynomial in θ, and its derivatives with

respect to θ are again polynomials in θ. Ellipticity therefore yields that, for any non-negative integer ℓ,

$$|\partial^\ell \big(\check{L}^H(\theta)^{-1}\big)_{jk}| \le C_\ell h^{\kappa'} |\theta|^{-\kappa-\ell}, \qquad (0 < |\theta| \le \frac{\pi}{2}) \tag{6.22}$$

where ∂^ℓ is any ℓ-order derivative with respect to $\theta = (\theta_1, \ldots, \theta_d)$, and κ and κ' are independent of ℓ. In fact we will use the much weaker condition

$$|\partial^\ell (\check{L}^H(\theta)^{-1} \check{I}_h^H(\theta) L^h(\theta))_{jk}| \le C_\ell h^{-\kappa'} |\theta|^{-\kappa-\ell} \qquad (h^{1-\sigma_*} \le |\theta| \le \frac{\pi}{2}) \tag{6.23}$$

where κ and κ' are independent of ℓ and $\sigma_* > 0$ sufficiently small. (The value $\sigma_* = 1/(2 \max_i m_{i*})$ will be shown suitable in the proof below.) This condition is much weaker because it would normally hold for non-elliptic systems as well, especially under assumption (6.4).

Thus, ellipticity is not explicitly used here. It is, however, related to the CGA and (6.13) assumptions, and to the size of λ (defined by (3.20)) that can be obtained. Real extensions to non-elliptic problems are discussed in Sec. .

Finally, we need to express in terms of symbols another assumption which in fact results form the CGA and (6.4) assumptions. The former implies that sufficiently smooth components are reduced as far as one wishes by means of the CGC step, while the latter, together with (6.13), imply that *harmonics* of sufficiently smooth components are practically unchanged by that step (since, by (6.2), (6.3) and (6.4), those harmonics contribute negligibly little to the right-hand side of the coarse-grid equations). Hence, the two assumptions together imply that, for $|\theta| \to 0$, the CGC action on the error components $E_\alpha \check{v}^B(\theta) e^{i\theta\alpha}$ (i.e., the component θ and its harmonics) practically gives $E_\alpha \check{T}(\theta) \check{v}^B(\theta) e^{i\theta\alpha}$, where $\check{T}(\theta)$ is the $qD \times qD$ matrix defined by

$$\begin{aligned} \check{T}(\theta)_{JK} &= \delta_{JK} T(\theta + \tau^{J-1}) Iq \\ T(\theta) = T(\theta_1, \ldots, \theta_d) &= \begin{cases} 0 & \text{if } -\frac{\pi}{2} \le \theta_j < \frac{\pi}{2} \text{ for } j = 1, \ldots, d \\ 1 & \text{otherwise.} \end{cases} \end{aligned} \tag{6.24}$$

Thus, as $|\theta| \to 0$, $M(\theta) \approx \tilde{M}_\nu(\theta) \underset{\text{def}}{=} \check{R}(\theta)^{\nu_1} \check{T}(\theta) \check{R}(\theta)^{\nu_2}$, and, in view of (3.20), we can make it our assumption that

$$\limsup_{\max(|\theta|,h)\to 0} \|\check{R}(\theta)^{\nu_2} \check{T}(\theta) \check{R}(\theta)^{\nu_1}\| \le \lambda. \tag{6.25}$$

Although actually resulting from other assumptions, (6.25) can separately be checked.

Incidentally, $\check{T}(\theta)$ represents the ideal performance of the CGC step: for sufficiently good I_h^H, A^H and I_H^h one would expect $M(\theta) \approx \tilde{M}_\nu(\theta)$ for all $|\theta| \le \pi/2$. In that case $\lambda' \approx \overline{\mu}_\nu^\nu$, where

$$\overline{\mu}_\nu = \sup_{|\theta| \le \pi/2} (\sigma(\tilde{M}_\nu(\theta)))^{1/\nu}. \tag{6.26}$$

This $\overline{\mu}_\nu$ is the familiar "*smoothing factor*" [G, §3.1], which can thus be used as a rough efficiency predictor. It is a very useful predictor because it is simpler to calculate than λ or λ' (especially in case relaxation is *strictly* consistently ordered, in which case $\overline{\mu}_\nu$ does not depend on ν and its calculation is reduced to computing $q \times q$ instead of $Dq \times Dq$ matrices) and, more important, it tells you what is the *ideal* performance one *can* obtain with a given relaxation scheme. Hence it allows a precise *separate* design of optimal relaxation. This design can usually further be reduced to the design of relaxation schemes for simple scalar equations (the factors of the principal determinant of the matrix operator; see [G84, §3.7]).

§7 Modified cycle and main theorem

As explained in Sec. 4, the bigrid cycle which we will actually analyze is a modification, involving negligible extra work, of the cycle (i)–(iii) defined in Sec. 2. To define the modified cycle we introduce a constant $0 < \sigma_1 < 1$ (whose value will further be specified later) and a small "distance from the boundary" $\rho = h^{\sigma_1}$, and for any $\rho' > 0$ we denote

$$\Omega_{\rho'} = \{x \mid x \in \Omega, |x - y| > \rho' \quad \text{for any} \quad y \notin \Omega\}. \tag{7.1}$$

$$\hat{\Omega}^h_{\rho'} = \Omega^h - \Omega^h_{\rho'}. \tag{7.2}$$

The steps of the modified cycles are the following. First, K_0 passes of Kaczmarz relaxation are performed in $\hat{\Omega}^h_{3\rho}$. Then steps (i), (ii) and (iii) of the unmodified cycle (Sec. 2) are carried out.

The number K_0 of boundary passes, to be specified below, will be independent of h, hence the total extra work involved will be at most $O(h^{\sigma_1})$ compared to the work in any of the other steps of the cycle. K_0 will in fact depend only on local properties of the fine-grid operator A^h (actually only on the quantity γ_1 defined in Theorem 5.1 and Remark 5.1) and on the accuracy ϵ to which we want to approximate the mode-analysis convergence factors by the modified-cycle convergence factor. This is the content of our main theorem.

Theorem 7.1. *Let λ be the mode-analysis convergence factor (3.20). Then, under the assumptions of Sec. 6, for any $\epsilon > 0$ there is $K_0 = K_0(\epsilon)$, independent of h, such that*

$$\frac{\|\overline{v}\|}{\|v\|} \le \lambda + \epsilon, \tag{7.3}$$

where v and $\overline{v}$ are the error functions (vectors) before and after the application of the modified cycle, and $\|\cdot\|$ is the ℓ_2 norm.

§8 Proof

8.1 Cutting away the boundary

Let $K_1 = (\gamma_1 \delta_2)^{-1}$ for some constant $\delta_2 = \delta_2(\epsilon)$ which will be selected below, where γ_1 is defined in Remark 5.1. Denote by $v^{[k]}$ the error vector obtained after k Kaczmarz boundary passes, and by $\|\tilde{A}v^{[k]}\|_{3\rho}$ the ℓ_2 norm of $\tilde{A}v^{[k]}$ confined to the relaxed domain $\hat{\Omega}^h_{3\rho}$, where $A = A^h$ and the normalized operator $\tilde{A}$ is defined in (6.1). Then, for some $1 \le k \le K_1$,

$$\|\tilde{A}v^{[k]}\|^2_{3\rho} \le \delta_2 \|v\|^2, \tag{8.1}$$

because otherwise, by Theorem 5.1, each Kaczmarz sweep would reduce the square error norm by more than $\gamma_1 \delta_2 \|v\|^2$, so the K_1 sweeps would reduce it below 0. Thus, taking $K_0 \le K_1$ to be that k for which $\|\tilde{A}v^{[k]}\|_{3\rho}$ is minimal, and denoting $v^{[K_0]}$ by v^a, we get

$$\|\tilde{A}v^a\|^2_{3\rho} \le \delta_2 \|v\|^2 \quad \text{and} \quad \|v^a\| \le \|v\|. \tag{8.2}$$

The second inequality is the result of the error-decreasing property of Kaczmarz relaxation (see Theorem 5.1).

We next note that, instead of calculating the final error $\overline{v}$ that results from applying Steps i) ii) and iii) starting with v^a as the initial error, it is enough to calculate the final error $\overline{v}^b$ resulting from an initial error

$$v^b = \varphi v^a, \tag{8.3}$$

where φ is a C^∞ function defined on $\mathbb{R}^d$ such that $0 \le \varphi \le 1$, $\varphi(x) = 0$ for $x \notin \Omega_\rho$, $\varphi(x) = 1$ for $x \in \Omega_{2\rho}$ and, for any x and y,

$$|\varphi(x) - \varphi(y)| \le C|x - y|/\rho. \tag{8.4}$$

Indeed, from (8.4) and Lemma 6.1 it follows that

$$\begin{aligned} \|\tilde{A}(v^b - v^a)\| &= \|\tilde{A}(1-\varphi)v^a\|_{3\rho} \\ &\le \|\tilde{A}v^a\|_{3\rho} + C\frac{h}{\rho}\|v^a\| \\ &\le C(\delta_2(\epsilon)^{1/2} + h^{1-\sigma_1})\|v\| \end{aligned} \tag{8.5}$$

the last inequality resulting from (8.2). Hence, for h sufficiently small and choosing $\delta_2(\epsilon) \le [\delta_1(\epsilon/2, 1)/(2C)]^2$, by (8.2) and Lemma 6.2 we get that

$$\|\overline{v} - \overline{v}^b\| \le \frac{\epsilon}{2}\|v\|. \tag{8.6}$$

The error function v^b, which vanishes outside Ω^h_ρ, can now be extended to the entire space $\mathbb{R}^d$, by defining $v^b(x) = 0$ for $x \notin \Omega^h$. This extended function has the Fourier decomposition, similar to (3.1),

$$v^b_\alpha = \int_{|\theta| \le \pi} \hat{v}^b(\theta) e^{i\theta\alpha}\, d\theta \tag{8.7}$$

where

$$\hat{v}^b(\theta) = (2\pi)^{-d} \sum_\beta v^b_\beta e^{-i\theta\beta}, \tag{8.8}$$

$\sum_\beta$ denoting summation over the entire grid of integer vectors $\beta = (\beta_1, \ldots, \beta_d)$. As implied by (8.8), $\hat{v}^b(\theta)$ is 2π-periodic (cf. (3.3)), and the range of integration meant in (8.7) is over one period, e.g., over the range (3.4). The *existence* of the Fourier transform (8.7) is trivial, since v^b has a bounded support and therefore $\hat{v}^b(\theta)$, defined by (8.8), is smooth and therefore also it is permitted, upon substituting (8.8) into the right-hand side of (8.7), to exchange the order of summation and integration. The integration then clearly vanishes unless $\alpha = \beta$, hence the summation yields the left-hand side of (8.7).

8.2 Separating away fringe components

Numerical processes which are fully local, such as Jacobi relaxation sweeps, will operate separately on each Fourier component (i.e., they will produce a new error function, whose new value for the Fourier amplitude $\hat{v}(\theta)$ will depend only on the old value of $\hat{v}(\theta)$ for the same θ). This is because such operators can directly be extended as constant-coefficient operators to the entire domain, at least for error functions vanishing near the boundaries. Other type of point-by-point relaxation sweeps, such as Gauss-Seidel and Kaczmarz, can also be so extended with negligible error (as we will see below), at least for error functions vanishing outside Ω^h_ρ where $\rho \gg h$. The inter-grid transfers I^H_h and I^h_H are truely local operators, but since they connect grid h with grid $2h$ they couple harmonics (see Sec. 3). The only truely non-local operation in the multigrid cycle is $(A^H)^{-1}$ — the solution, in Step (ii), of the coarse-grid equations. The main idea of our proof is the observation that even $(A^H)^{-1}$ can be regarded as a local operation, provided it is confined to error functions vanishing outside Ω^h_ρ (except for possible residues smaller than any power of h) and having non-vanishing Fourier components $\hat{v}(\theta)$ only in the range $|\theta| \gg h/\rho$. Since $\rho = h^{\sigma_1} \gg h$, this range includes all components except for very smooth ones. Our next step in the proof is therefore to separate from v^b those very smooth components, for which special estimates (using their smoothness rather than mode analysis) will then be applied.

The general form of v^0, the very smooth part to be separated from v^b, will be

$$v^0_\alpha = \int_{|\theta|\le\pi} \psi(\theta)\hat{v}^b(\theta)e^{i\theta\alpha}\, d\theta \tag{8.9}$$

where ψ is a 2π-periodic function (cf. (3.3)) such that

$$\psi(\theta) = 1 \quad \text{for} \quad |\theta| \le \eta_0 \quad \text{and} \quad \psi(\theta) = 0 \quad \text{for} \quad \eta_1 \le |\theta| \le \pi \tag{8.10}$$

and where $h^{1-\sigma_1} \ll \eta_0 < \eta_1$. The choice of η_0, η_1 and other properties of ψ will be described below. Since in the bigrid mode analysis each component θ is coupled to

its "harmonics" $\theta^j = \theta + \tau^j \pmod{2\pi}$ ($j = 1, \ldots, D-1$; see Sec. 3), we should also separate from v^b the harmonics of the very smooth components, i.e., the functions

$$v^j_\alpha = \int \psi(\theta + \tau^j)\hat{v}^b(\theta)e^{i\theta\alpha}\, d\theta, \qquad (j = 1, \ldots, D-1). \tag{8.11}$$

On the coarse grid these functions become very smooth, so their analysis, as we will see, will be a combination of mode analysis (in relaxation) and the use of smoothness-dependent estimates (in the coarse grid correction).

The remaining function

$$v^c = v^b - \sum_{j=0}^{D-1} v^j, \tag{8.12}$$

which represents the bulk of the error, will be analyzed by mode analysis. To do this, however, we need that v^c, like v^b, (very nearly) vanishes in a sufficiently large neighborhood of the boundary (so that the deviation of the multigrid processes from having constant coefficient over the entire space will have negligible effect). This is obtained by choosing sufficiently smooth ψ. Specifically, we choose $\eta_1 = \eta_0 + \eta_2$ with

$$\eta_2 = O(h^{1-\sigma_2}), \qquad \eta_0 = O(h^{1-\sigma_3}), \qquad \text{and} \quad 1 > \sigma_3 > \sigma_2 > \sigma_1, \tag{8.13}$$

and then construct $\psi(\theta) \in C^\infty$ such that

$$0 \le \psi(\theta) \le 1 \qquad \text{and} \qquad |\partial^\ell \psi(\theta)| \le C_\ell \eta_2^{-\ell} \tag{8.14}$$

for any θ and any ℓ-order derivative ∂^ℓ, ($\ell = 0, 1, 2, \ldots$; the exact values of η_2, η_0, σ_2 and σ_3 will be further specified later). As we will see below this entails that v^c practically vanishes outside $\Omega^h_{\rho/2}$.

Since all our cycle processes are linear we can apply them separately to v^c and to $v^d = \sum_{j=0}^{D-1} v^j$ and estimate their separate contributions to the final error $\overline{v}$. We denote these contributions by $\overline{v}^c$ and $\overline{v}^d$, respectively.

8.3 Estimating the main error $\overline{v}^c$

By (8.12), (8.9), (8.11) and similarly to (3.15), the main error v^c can be decomposed as

$$v^c_\alpha = \int^0 \psi_0(\theta) E_\alpha \check{v}^b(\theta) e^{i\alpha\theta}\, d\theta \tag{8.15}$$

where $\psi_0(\theta) = 1 - \psi(\theta)$ and $\check{v}^b(\theta)$ is the vector

$$\check{v}^b(\theta) = (\hat{v}^b(\theta^0)^\dagger, \ldots, \hat{v}^b(\theta^{D-1})^\dagger)^\dagger. \tag{8.16}$$

Applying the bigrid processes to v^c *in the infinite domain* would successively produce the functions v^{c1}, $v^{c2},\dots,v^{c6}$ defined by

$$v_\alpha^{ct}=\int^0 \psi_0(\theta)E_\alpha^t M_t(\theta)\check{v}^b(\theta)e^{i\theta\alpha}\,d\theta,\qquad (t=1,\dots,6) \tag{8.17}$$

where $M_1(\theta)=\check{R}(\theta)^{\nu_1}$, $M_2(\theta)=\check{L}^h(\theta)M_1(\theta)$, $M_3(\theta)=\check{I}_h^H(\theta)M_2(\theta)$, $M_4(\theta)=\check{L}^H(\theta)^{-1}M_3(\theta)$, $M_5(\theta)=\check{I}_H^h(\theta)M_4(\theta)$ and $M_6(\theta)=\check{R}(\theta)^{\nu_2}[M_1(\theta)-M_5(\theta)]=M(\theta)$ (cf. (3.19)) and $E_\alpha^1=E_\alpha^2=E_\alpha^5=E_\alpha^6=E_\alpha$ while $E_\alpha^3=E_\alpha^4=I_q$. We have to prove that, with negligible differences, the same functions will be produced by applying the corresponding *real* processes, in the *real* domain Ω^h with its *real* discrete boundary conditions. Since the boundary conditions for the error functions are always homogeneous, it is enough to prove that each v^{ct} practically vanishes outside $\Omega^h_{\rho/2}$. Indeed, its j-th component, $v^{ct,j}$, is a function which, by (8.17), has the general form

$$v_\alpha^{ct,j}=\sum_\beta\sum_{k=1}^q v_\beta^{b,k}\int_{|\theta|\le\pi}\psi_*(\theta)M_{tjk}(\theta,h)e^{i\theta(\alpha-\beta)}\,d\theta,\quad (j=1,\dots,q) \tag{8.18}$$

where $\psi_*(\theta)=1-\psi(\theta^0)$, θ^0 denoting, here and below in this section, the lowest harmonic of θ, i.e., $-\frac{\pi}{2}\le\theta_i^0<\frac{\pi}{2}$ and $(\theta_i-\theta_i^0)/\pi$ is an integer $(i=1,\dots,d)$. By (8.14) and (8.10),

$$\begin{aligned} &0\le\psi_*(\theta)\le 1,\quad |\partial^\ell\psi_*(\theta)|\le C_\ell\eta_2^{-\ell}\\ &\text{and}\quad \psi_*(\theta)=0\quad\text{for}\quad |\theta^0|\le\eta_0. \end{aligned} \tag{8.19}$$

By the assumptions in Sec. 6.5, each $M_{tjk}(\theta,h)$ is a rational function of trigonometric functions of θ, unbounded only at $\tau^0,\dots,\tau^{D-1}$, satisfying

$$|\partial^\ell M_{tjk}(\theta,h)|\le C_\ell h^{-\kappa'}|\theta^0|^{-\kappa-\ell},\quad (h^{1-\sigma_*}\le|\theta^0|\le\frac{\pi}{2}) \tag{8.20}$$

where ∂^ℓ is any ℓ-order (partial) derivative with respect to $\theta=(\theta_1,\dots,\theta_d)$, and κ and κ' are independent of ℓ. Choosing for each β an index $1\le j_\beta\le d$ such that $|\alpha_{j_\beta}-\beta_{j_\beta}|=|\alpha-\beta|\underset{\text{def}}{=}\max_i|\alpha_i-\beta_i|$ and then integrating by parts ℓ times with respect to θ_{j_β} the integral in (8.18), one obtains

$$\begin{aligned} |v_\alpha^{ct,j}|&\le C\sum_{\beta,k}|v_\beta^{b,k}|\;|\alpha-\beta|^{-\ell}\max_{|\theta|\le\pi}\left|\frac{\partial^\ell}{\partial\theta_{j_\beta}^\ell}\left[\psi_*(\theta)M_{tjk}(\theta,h)\right]\right|\\ &\le C_\ell\sum_{\beta,k}|v_\beta^{b,k}|\;|\alpha-\beta|^{-\ell}\sum_{0\le\ell'\le\ell}\eta_2^{-\ell'}h^{-\kappa'}\eta_0^{-\kappa-(\ell-\ell')} \end{aligned}$$

the second inequality resulting from (8.19), (8.20) and from assuming $\sigma_*<\sigma_3$ (hence $h^{1-\sigma_*}<\eta_0$, hence the applicability of (8.20) for any $|\theta^0|>\eta_0$). Now,

for any $\alpha \notin \Omega^h_{\rho/2}$ and any β such that $v^{b,k}_\beta \neq 0$ (hence $\beta \in \Omega^h_\rho$) we clearly have $|\alpha h - \beta h| \geq \rho/2 = h^{\sigma_1}/2$. Hence, for $\ell > d/2$, applying Cauchy-Schwarz inequality and using magnitudes (8.13) for η_2 and η_0, the inequality above yields

$$\begin{aligned} |v^{ct,j}_\alpha| &\leq C_\ell \|v^b\| \max_{0 \leq \ell' \leq \ell} h^{(\sigma_1 - 1)d/2 - \kappa' - (1-\sigma_3)\kappa + (\sigma_2 - \sigma_1)(\ell - \ell') + (\sigma_2 - \sigma_1)\ell'} \\ &\leq C_\ell h^{-\kappa'' + (\sigma_2 - \sigma_1)\ell} \|v^b\| \end{aligned}$$

where κ'' is independent of ℓ.

Thus, for $\alpha \notin \Omega^h_{\rho/2}$, since ℓ is arbitrary, we see that $|v^{ct,j}_\alpha|$ is smaller than any power of h. Hence, each v^{ct} satisfies the homogeneous boundary conditions with an error smaller than any power of h. Due to the stability of the bigrid processes (6.10)–(6.14), we can therefore conclude that, for any desired ℓ,

$$\|v^{c6} - \overline{v}^c\| \leq C_\ell h^\ell \|v^b\|. \tag{8.21}$$

In addition, by (3.19)–(3.20) and the Parseval identity,

$$\|v^{c6}\| \leq \lambda \|v^c\|, \tag{8.22}$$

which together with (8.21) yields the required estimate for $\overline{v}^c$.

8.4 Estimating fringe error $\overline{v}^d$

Applying step (i) (ν_1 relaxation sweeps) to v^d one obtains v^e, defined by

$$v^e_\alpha = \int^0 \psi(\theta) E_\alpha \check{R}(\theta)^{\nu_1} \check{v}^b(\theta) e^{i\theta\alpha}\, d\theta, \tag{8.23}$$

and, as before, it is immaterial whether the relaxation incorporates the boundary conditions or not, since v^e practically vanishes outside $\Omega^h_{\rho/2}$. Observing that $\tau^0 = 0$ in (3.16) and writing $E_\alpha = E + E'_\alpha$, where $E = (I_q, 0, \ldots, 0)$, the smooth part of v^e is v^f, defined by

$$v^f_\alpha = \int^0 \psi(\theta) E \check{R}(\theta)^{\nu_1} \check{v}^b(\theta) e^{i\theta\alpha}\, d\theta. \tag{8.24}$$

Since v^f practically vanishes near the boundary and its transform, by (8.24) and (8.10), contains Fourier components $e^{i\theta\alpha}$ only in the range $|\theta| \leq \eta_0 = h^{1-\sigma_3}$, our assumption (6.8) implies that, for h sufficiently small and any desired constant $\epsilon_0 > 0$,

$$\|\tilde{A}^h v^f\| \leq \delta(\epsilon_0) \|v^f\|. \tag{8.25}$$

Hence, by the CGA assumption, applying CGC (the coarse grid correction step) to v^f will give contribution smaller than $\epsilon_0\|v^f\|$. Since clearly $\|v^f\| \le \|v^e\| \le C\|v^b\| \le C\|v\|$ (the second inequality resulting from (6.20)), by (6.10) the contribution $\overline{v}^f$ of v^f to $\overline{v}^d$ satisfies

$$\|\overline{v}^f\| \le C\epsilon_0\|v\| \le \frac{\epsilon}{12}\|v\|, \tag{8.26}$$

where the second inequality is obtained by a suitable choice of ϵ_0.

The non-smooth part of v^e is $v^g = v^e - v^f$, given by

$$\begin{aligned} v^g_\alpha &= \int^0 \psi(\theta) E'_\alpha \check{R}(\theta)^{\nu_1} \check{v}^b(\theta) e^{i\theta\alpha}\, d\theta \\ &= \int^0 \psi(\theta) E_\alpha \check{T}(\theta) \check{R}(\theta)^{\nu_1} \check{v}^b(\theta) e^{i\theta\alpha}\, d\theta. \end{aligned} \tag{8.27}$$

(see definition of $\check{T}(\theta)$ in (6.24)). Since the only non-vanishing Fourier components in (8.27) are of the form $\exp i(\theta + \tau^j)$, with $|\theta| \le \eta_0$ and $j \ge 1$, it follows from (6.2) and (6.3) that

$$\begin{aligned} \|(I_h^H L^h v^g)_i\| &\le C\eta_0^{m_i} h^{-m_{i*}} \|v^g\| \\ &\le C h^{1-\sigma_3 m_i} \|v^g\|, \end{aligned} \tag{8.28}$$

the second inequality resulting from (8.13) and (6.4). Hence, by (6.20)

$$\|I_h^H L^h v^g\| \le C h^{1-\sigma_3 m_*} \|v^b\|, \tag{8.29}$$

where $m_* = \max_i m_i$. Since, for reasons as before, $I_h^H L^h v^g$ practically vanishes near the boundary (being smaller than any power of h in $\Omega^h_{\rho/2}$), it would be obtained whether A^h and I_h^H are applied with or without the real boundary conditions ($A^h = L^h$ in the later case). We now apply to it the *real* (including boundary conditions) $(A^H)^{-1}$. Using (6.13), (6.14) and (6.10) we conclude from (8.29) that

$$\|R^{\nu_2} I_H^h (A^H)^{-1} I_h^H L^h v^g\| \le C h^{1-\sigma_3 m_*} \|v^b\|. \tag{8.30}$$

Choosing $\sigma_3 < 1/m_*$ and a sufficiently small h, the right-hand side of (8.30) is smaller than $\frac{\epsilon}{12}\|v\|$ and hence, by (8.26),

$$\|\overline{v}^d - R^{\nu_2} v^g\| \le \frac{\epsilon}{6}\|v^b\|. \tag{8.31}$$

By (8.27), and whether boundary conditions are used in R or not, we have

$$R^{\nu_2} v^g = \int^0 \psi(\theta) E_\alpha \check{R}(\theta)^{\nu_2} \check{T}(\theta) \check{R}(\theta)^{\nu_1} \check{v}^b(\theta) e^{i\theta\alpha}\, d\theta.$$

Hence, using (8.14), (6.25) and the Parseval identity,

$$\|R^{\nu_2} v^g\| \le \lambda \|v^d\|, \tag{8.32}$$

yielding, together with (8.31), the required bound on $\overline{v}^d$.

8.5 Summary

The Fourier components of v^{c6} are only in the range $\pi/2 \geq |\theta| \geq \eta_1 = \eta_0 + \eta_2$ and its harmonics, while those of $R^{\nu_2} v^g$ are in the range $|\theta| \leq \eta_0$ and its harmonics. The only overlap (in terms of lowest harmonics) is thus in the range $\eta_0 \leq |\theta| \leq \eta_1$, whose volume is less than $O(\eta_2/\eta_0) = O(h^{\sigma_3 - \sigma_2})$ compared to the volume of either $\{|\theta| \leq \eta_0\}$ or $\{|\theta| \geq \eta_1\}$. Hence, writing $\eta_0 = \eta_* h^{1-\sigma_3}$, we can choose $1 \leq \eta_* \leq 2$ such that the components in the overlap contribute at most $O(h^{\sigma_3 - \sigma_2})$ to either $\|v^{c6}\|^2$ or $\|R^{\nu_2} v^g\|^2$, so that

$$\|v^{c6} + R^{\nu_2} v^g\|^2 \leq (\|v^{c6}\|^2 + \|R^{\nu_2} v^g\|^2)(1 + Ch^{\sigma_3 - \sigma_2}).$$

Using (8.21) and (8.31) we hence have, for sufficiently small h,

$$\begin{aligned}
\|\overline{v}^b\|^2 &= \|\overline{v}^c + \overline{v}^d\|^2 \\
&\leq (\|v^{c6} + R^{\nu_2} v^g\| + \frac{\epsilon}{5}\|v^b\|)^2 \\
&\leq (\|v^{c6}\|^2 + \|R^{\nu_2} v^g\|^2)(1 + Ch^{\sigma_3 - \sigma_2}) \\
&+ 2\frac{\epsilon}{5}\|v^b\|(\|v^{c6}\| + \|R^{\nu_2} v^g\|) + \frac{\epsilon^2}{25}\|v^b\|^2
\end{aligned}$$

Hence, by (8.22) and (8.32)

$$\|\overline{v}^b\|^2 \leq \lambda^2(\|v^c\|^2 + \|v^d\|^2)(1 + Ch^{\sigma_3 - \sigma_2}) + \frac{2\epsilon}{5}\|v^b\|\lambda(\|v^c\| + \|v^d\|) + \frac{\epsilon^2}{25}\|v^b\|^2. \quad (8.33)$$

Furthermore, since in (8.9) and (8.11) $0 \leq \psi \leq 1$, the Parseval identity yields

$$\|v^b\|^2 = \|v^c + v^d\|^2 \geq \|v^c\|^2 + \|v^d\|^2, \quad (8.34)$$

and therefore

$$\|v^c\| \leq \|v^b\|, \quad \|v^d\| \leq \|v^b\|. \quad (8.35)$$

For sufficiently small h, by (8.33), (8.34) and (8.35)

$$\begin{aligned}
\|\overline{v}^b\| &\leq \lambda^2\|v^b\|^2(1 + Ch^{\sigma_3 - \sigma_2}) + (\frac{4\epsilon}{5}\lambda + \frac{\epsilon^2}{25})\|v^b\|^2 \\
&\leq (\lambda + \frac{\epsilon}{2})^2\|v^b\|^2 \quad (8.36) \\
&\leq (\lambda + \frac{\epsilon}{2})^2\|v\|^2,
\end{aligned}$$

the last inequality resulting from (8.3) and (8.2). Thus, using (8.6), we obtain (7.3). ∎

§9 Extension to variable coefficients

Using essentially the same proof, various modifications and extensions can be introduced to the above theory, strengthening Theorem 7.1 in various directions. We discuss below several of the important extensions, one at a time. The various technique we show for modifying the theory can then also be used *in combination with each other*, to prove any combination of the discussed extensions.

Extensions to variable coefficients are discussed in more detail in this section, others are described in Sec. 10.

9.1 Zero-volume discontinuities

The above theory allows the coefficients of the discrete operator A^h to have arbitrary behavior, including discontinuities, in some neighborhood of the boundary, provided that that neighborhood tends to have zero relative volume as the meshsize h tends to zero: e.g., the width of the neighborhood can be $O(h^{\binom{\sigma_1}{h}})$, for any $\sigma_1 > 0$. Actually, the proof did not use the fact that that neighborhood was just a neighborhood of the boundary. Thus, *the coefficients can change arbitrarily, including discontinuities, provided this happens in a region whose relative volume tends to zero*—and provided, of course, that the CGA assumption still holds. The local mode analysis is then simply applied separately in each subdomain where A^h has constant coefficients. When the coefficients are *strongly* discontinuous (jumping in their order of magnitude), the satisfaction of the CGA assumption is not trivial; it is indeed obtained only when special coarsening techniques are used (like those described in [ABDP] or [G, §4.4]; see related remarks in Sec. below).

9.2 Smooth coefficients

The *interior* discrete operator L^h (i.e., the operator outside those vanishing neighborhoods where the behavior of the coefficients can be arbitrary) was assumed to have constant coefficients (in each subdomain, the subdomains being separated from each other by interfaces with vanishing total volume). In fact, it is enough to assume that the coefficients are *sufficiently smooth* (in each subdomain). To show this, we first have to reinterpret the local mode analysis itself by freezing coefficients (as in [BD, p. 98] or [G, §4]). That is, the mode analysis of Sec. 3 above is first applied to $L^h(x)$, the discrete operator in the infinite space whose constant coefficients are those of L^h at the point x. In this way a convergence factor $\lambda(x)$, defined by (3.20), is calculated for every x. From this

$$\lambda = \sup_{x \in \Omega_\rho} \lambda(x) \tag{9.1}$$

is determined (in practice by a simple maximization routine). This λ is the *mode analysis predicted convergence factor*, which should be used in Theorem 7.1 for the variable-coefficients case.

To extend to this situation the proof in Sec. 8, we introduce a Gårding-type partition of unity on scale $s_p = h^{\sigma_p}$, where

$$0 < \sigma_1 < \sigma_p < \sigma_2 < \sigma_3 \tag{9.2a}$$

(subscript p for "partition"). That is, we introduce a set of scalar continuum functions $\varphi^{h;k}(x)$, $(k = 1, 2, \ldots, K^h)$, satisfying for any $x \in \Omega$ and $1 \le k \le K^h$ the following conditions:

$$0 \le \varphi^{h;k}(x) \le 1 \tag{9.2b}$$

$$\sum_{k=1}^{K^h} [\varphi^{h;k}(x)]^2 = 1 \tag{9.2c}$$

$$\varphi^{h;k}(x) = 0 \quad \text{for} \quad |x - t_k^h| \ge s_p \tag{9.2d}$$

$$|(\frac{\partial}{\partial x_1})^{\ell_1} \cdots (\frac{\partial}{\partial x_d})^{\ell_d} \varphi^{h;k}(x)| \le C_{\ell_1+\cdots+\ell_d} s_p^{-(\ell_1+\cdots+\ell_d)} \tag{9.2e}$$

the last inequality holding for any vector of non-negative integers $(\ell_1, \ldots, \ell_d)$. For each k we then apply to $v^{b;k} = \varphi^{h;k} v^b$ the same proof steps formerly applied to v^b, defining $v^{0;k}$, $v^{j;k}$, $v^{c;k}$, $v^{ct;k}$, $v^{ct;k,j}$, $\overline{v}^{c;k}$, $v^{d;k}$, $v^{e;k}$, $v^{f;k}$, $v^{g;k}$, $\overline{v}^{d;k}$ and $\overline{v}^{b;k} = \overline{v}^{c;k} + \overline{v}^{d;k}$ analogously to the definition of the corresponding functions in Secs. 8.2–8.4. Each of these functions is produced from a previous one by a certain operator, for which the following 3 claims should be proved.

Claim I. Applying the real (possibly variable-coefficient) operator produces, up to a negligible error, the same function as applying the corresponding frozen-coefficient operator, i.e., the infinite-domain constant-coefficients operator whose coefficients coincide with the real ones at t_k^h.

Claim II. The support of the function is practically unchanged by the operator, i.e., if $\sigma_3 - \sigma_p > \sigma_{3p} > 0$ then, for

$$|x - t_k^h| \ge s_p(1 + h^{\sigma_{3p}}) \tag{9.3}$$

the value of the function at x is smaller than any power of h.

Claim III. Except for negligible errors $v^{0;k} = \varphi^{h;k} v^0$, $v^{c;k} = \varphi^{h;k} v^c$, $v^{ct;k} = \varphi^{h;k} v^{ct}$, etc.

It is enough to discuss here how these three claims are proved in the step of applying $(A^H)^{-1}$ (the step of producing $v^{c4;k}$), because all other steps are

simpler and require less smoothness of coefficients. In terms of this example we will also describe the claims in more concrete terms, specify what is meant by "negligible errors" and "practically unchanged". Denote by L_k^H the constant-coefficient operator $L^H(t_k^h)$. To prove Claim I we have to show, for any $\epsilon_* > 0$, that, for sufficiently small $H = 2h$,

$$\|((A^H)^{-1} - (L_k^H)^{-1})v^{c3;k}\| \le \epsilon_* \|(L_k^H)^{-1} v^{c3;k}\|. \tag{9.4}$$

Since we assume that Claim I has already been proved for all previous steps, we know that $v^{c3;k}$ is made of Fourier components $e^{i\theta\cdot x/H}$ only in the range $2\eta_0 \le \theta \le \pi$ (the factor 2 appearing because we are dealing here with the coarse grid H, and hence it is $H = 2h$ that we use in the exponent denominator). As in Sec. 8.3 it can therefore be shown that, with negligible error,

$$(L_k^H)^{-1} v^{c3;k} = \int^0 \psi_0(\theta) M_4^k(\theta) \check{v}^{b;k}(\theta) e^{i\theta\alpha} d\theta, \tag{9.5}$$

where $M_4^k(\theta)$ is the symbol $M_4(\theta)$ for coefficients frozen at t_k^h; and where it does not matter whether L_k^H employs the real boundary conditions or describes the infinite-domain operator. Thus, instead of (9.4) it is enough to prove the following condition.

Coefficient Condition 9.1. *For any $\epsilon_* > 0$ and $0 < \sigma_3 < 1$ there exist $H_* > 0$, $0 < \sigma_p < \sigma_3$ and a positive integer ℓ such that, for all $0 < H < H_*$, the inequality*

$$\|(A^H)^{-1}(L_k^H - A^H)v^H\| \le \epsilon_* \|v^H\| \tag{9.6}$$

holds for any v^H having the Fourier representation

$$v_\alpha^H = v^H(\alpha H) = \int_{h^{1-\sigma_3} \le |\theta| \le \pi} \hat{v}^H(\theta) e^{i\theta\alpha} d\theta \tag{9.7}$$

and satisfying

$$\|v^H(x)\| \le C_\ell H^\ell \quad \text{for all} \quad |x - t_k^h| \ge C H^{\sigma_p}. \tag{9.8}$$

This is a qualitative coefficient-smoothness condition that we can most simple *assume* to hold. But we can also make the following observations about its translation into more familiar coefficient smoothness requirements.

If A^H is uniformly continuous, it is clear that $\|(L_k^H - A^H)v^H\|$ can be made arbitrarily small (compared to $\|A^H v^H\|$) by choosing sufficiently small H, since the subdomain (9.3) can be made arbitrarily small. Unfortunately, if $L_k^H - A^H$ is not smooth enough, then, for some particular functions v^H, the function $(L_k^H - A^H)v^H$

may turn out much smoother than $A^H v^H$, hence the former may be amplified much more than the latter upon being operated on by $(A^H)^{-1}$. This is not a mere theoretical trouble: it can actually happen, and in such a situation the local mode analysis predictions can indeed be inaccurate. (Hence the above Coefficient Condition 9.1, or at least some weaker form of it, is a necessary condition). Thus, to satisfy (9.6) it is necessary to assume enough smoothness of A^H.

For example, if A^H is uniformly elliptic and its coefficients have uniformly bounded derivatives (or difference-quotients) of sufficiently high orders on scale s_A (i.e., s_A^ℓ times any ℓ-order derivative of any coefficient is uniformly bounded, for all $\ell \leq \ell_{\max}$, $\ell_{\max}$ being sufficiently large), where $s_A = h^{\sigma_A}$ and $\sigma_p > \sigma_A \geq 0$, then (9.6) can be proved by infinite-space Fourier Analysis. The proof is given in Sec. 9.2.1 below. (Notice that it indeed contains a proof of Claim II, too.)

The smoothness requirements on A^H can be reduced by employing higher Sobolev norms instead of the L_2 norm used here. Such norms give less weights to smooth components (cf. Sec. below). In particular, if a high enough norm is used, then the fact that $(L_k^H - A^H)v^H$ is smoother than $A^H v^H$ makes its norm correspondingly smaller, too, and consequently the trouble described above disappears. So, with a suitable norm, (9.6) may be proved for any A^H which is uniformly continuous on scale s_A.

Once Claim I has been established for $(A^H)^{-1}$, Claim II can easily also be proved, by applying the proof in Sec. 8.3, with $M_4^k(\theta)$ replacing $M_4(\theta)$ and σ_p replacing σ_1. Claim III is also easily established since, following Claim I, it needs only be proved for the constant coefficient operator in the infinite space, and only for v^H of the form (9.7), where, by (9.2a) the wavelengths in v^H are much shorter than s_p, the scale on which $\varphi^{h;k}$ is infinitely smooth. The proof is given in Sec. 9.2.2 below.

That proof, written for the operator $(L_k^H)^{-1}$, is easily adapted to simpler operators (for which the restriction on the Fourier transform support, appearing in (9.7) as a restriction of the integration domain, neither always holds nor is needed); such as the operator (8.9) of separating away the smooth part v^0 from the error v^b, which now analogously separates $v^{0;k}$ from $v^{b;k}$. (The proof here uses (8.14) for $\ell = 1$, (8.13) and the relation $\sigma_2 > \sigma_p$). Thus we get $v^{0;k} = \varphi^{h;k} v^0$, and similarly $v^{j;k} = \varphi^{h;k} v^j$ $(j = 1, \ldots, D-1)$. We then prove Claims I, II and III subsequently for $v^{ct;k}$ $(t = 1, \ldots, 6)$ and for $R^{\nu_1} v^{b;k}$, $v^{e;k}$, $v^{f;k}$, $v^{g;k}$ and $R^{\nu_2} v^{g;k}$, in the same manner demonstrated above for $v^{c4;k}$. Claim I implies that for each of these we can obtain estimates analogous to those in Secs. 8.3–8.5. These estimates (together, in fact, with estimates in Claim I itself, such as (9.4)) thus yield, analogously to (8.36),

$$\|\overline{v}^{b;k}\|^2 \leq \left(\lambda\left(t_k^h\right) + \frac{\epsilon}{2}\right)^2 \|v^{b;k}\|^2. \tag{9.9}$$

Using this together with (9.1), (9.2c) and Claim III (applied to $\overline{v}^{b;k}$), it follows

that (upto negligible errors)

$$\begin{aligned}
\|\overline{v}^b\|^2 &= \sum_\alpha |\overline{v}^b_\alpha|^2 \\
&= \sum_\alpha \sum_k [\varphi^{h;k}(\alpha h)]^2 |\overline{v}^b_\alpha|^2 \\
&= \sum_\alpha \sum_k |\overline{v}^{b;k}_\alpha|^2 \\
&= \sum_k \|\overline{v}^{b;k}\|^2 \\
&\le (\lambda + \frac{\epsilon}{2})^2 \sum_k \|v^{b;k}\|^2 \\
&= (\lambda + \frac{\epsilon}{2})^2 \sum_k \sum_\alpha [\varphi^{h;k}(\alpha h)]^2 |v^b_\alpha|^2 \\
&= (\lambda + \frac{\epsilon}{2})^2 \|v^b\|^2 \\
&\le (\lambda + \frac{\epsilon}{2})^2 \|v\|^2, \qquad (9.10)
\end{aligned}$$

the last inequality resulting from (8.3) and (8.2). Since the derivation of (8.6) did not depend on A^h having constant coefficients, we can use it here, which together with (9.10) yields (7.3).

9.2.1 Proof of Coefficient Condition 9.1. As stated above, we want to prove (9.6) under (9.7)–(9.8), assuming A^H to be uniformly elliptic and sufficiently smooth on scale s_A, where

$$s_A = h^{\sigma_A} \quad \text{and} \quad \sigma_3 > \sigma_p > \sigma_A \ge 0. \qquad (9.11)$$

We will first write the ellipticity and smoothness assumptions in terms of the *symbol* of A^H, which is the $q \times q$ matrix $\hat{a}^H_\alpha(\theta)$ defined by

$$(A^H \hat{v}^H e^{i\theta\alpha})_{\alpha'} = \hat{a}^H_{\alpha'}(\theta) \hat{v}^H e^{i\theta\alpha'} \qquad (9.12)$$

holding for any constant q-vector $\hat{v}^H$. Thus, each term in $\hat{a}^H_\alpha(\theta)$ is a trigonometric polynomial in θ with coefficients depending on the grid point α, reflecting the coefficients of A^H at α.

Without loss of generality we can assume, by simple translation, that $t^h_k = 0$, so that

$$(\mathrm{L}^H_k \hat{v}^H e^{i\theta\alpha})_{\alpha'} = \hat{a}^H_0(\theta) \hat{v}^H e^{i\theta\alpha'}. \qquad (9.13)$$

With the support of v^H_α hence being essentially $|\alpha H| \le C H^{\sigma_p}$, it then follows from the Fourier inversion

$$\hat{v}^H(\theta) = (2\pi)^{-d} \sum_\alpha v^H_\alpha e^{-i\theta\alpha} \qquad (9.14)$$

that $\hat{v}^H(\theta)$ is a very smooth function, and in particular, using the Parseval identity, that

$$[\int_{|\theta|\leq\pi} \|\hat{v}^H(\theta) - \hat{v}^H(\theta-\eta)\|^2 d\theta]^{1/2} \leq C|\eta|H^{\sigma_p-1}\|v^H\|. \tag{9.15}$$

Also, without changing the smoothness-on-scale-s_A of A^H, we can then assume

$$\hat{a}^H_\alpha(\theta) = 0 \quad \text{for} \quad |\alpha H| \geq Cs_A \tag{9.16}$$

(since the support of v^H_α is far smaller than that). Hence $hata^H_\alpha(\theta)$ as a function of α, has the Fourier expansion

$$\hat{a}^H_\alpha(\theta) = \int_{|\eta|\leq\pi} \hat{a}^H(\eta,\theta)e^{i\eta\alpha}d\eta, \tag{9.17}$$

where

$$\hat{a}^H(\eta,\theta) = (2\pi)^{-d}\sum_\alpha \hat{a}^H_\alpha(\theta)e^{-i\eta\alpha}. \tag{9.18}$$

The assumption that A^H is uniformly elliptic means that there is a polynomial P_A such that $P_A(\omega) > 0$ for any $\omega \neq 0$, for which

$$\|\hat{a}^H_\alpha(\theta)^{-1}\| \leq \left[P_A\left(\frac{|\theta|_*}{H}\right)\right]^{-1} \tag{9.19}$$

and

$$\|\hat{a}^H_\alpha(\theta)\| \leq CP_A\left(\frac{|\theta|_*}{H}\right), \tag{9.20}$$

where we have introduced, for any $\theta = (\theta_1,\ldots,\theta_d)$, the notation

$$|\theta|_* = \max_{1\leq j\leq d}|\theta_j|_*, \qquad |\theta_j|_* = \min_{integer\ \kappa}|\theta_j - 2\pi\kappa|,$$

and where $\|\cdot\|$ is the ℓ_2 matrix operator norm. Furthermore, since $\hat{a}^H_\alpha(\theta)$ is a smooth function (a trigonometric polynomial, in fact) in θ, we can also assume, for any $|\eta|_* \leq \frac{1}{C}|\theta|_*$, that

$$\|[\hat{a}^H_\alpha(\theta) - \hat{a}^H_\alpha(\theta-\eta)]\| \leq C\frac{|\eta|_*}{|\theta|_*}P_A\left(\frac{|\theta|_*}{H}\right).$$

By (9.18) and (9.16), this can also be expressed as

$$\|[\hat{a}^H(\eta,\theta) - \hat{a}^H(\eta,\theta-\eta)]\| \leq C\frac{|\eta|_*}{|\theta|_*}P_A\left(\frac{|\theta|_*}{H}\right)\left(\frac{s_A}{H}\right)^d,$$

or, by (9.19),

$$\|\hat{a}_0^H(\theta)^{-1}[\hat{a}^H(\eta,\theta)-\hat{a}(\eta,\theta-\eta)]\| \le C\frac{|\eta|_*}{|\theta|_*}\left(\frac{s_A}{H}\right)^d. \tag{9.21}$$

Similarly, from (9.20) and (9.19), together with (9.18) and (9.16), for $|\theta'|_* \le C|\theta|_*$

$$\|\hat{a}_0^H(\theta)^{-1}\hat{a}^H(\eta,\theta')\| \le C\left(\frac{s_A}{H}\right)^d. \tag{9.22}$$

The assumption that A^H is smooth of order ℓ on scale s_A can be written as

$$\|\partial_H^\ell \hat{a}_\alpha^H(\theta)\| \le CP_A\left(\frac{|\theta|_*}{H}\right)s_A^{-\ell} \tag{9.23}$$

for any $0 \le \ell \le \ell_{\max}$, where ∂_H^ℓ is any ℓ-order difference-quotient operator on the grid $\alpha = x/H$. Let $\eta = (\eta_1,\dots,\eta_d)$ and let $j = j_\eta$ be such that $|\eta_j|_* = |\eta|_*$. The summation in (9.18) includes a summation over α_j. Since $\hat{a}_\alpha^H(\theta)$ has a bounded support, we can perform ℓ summations by parts on that α_j summation, obtaining, by (9.16) and (9.23), the estimate

$$\begin{aligned}\|\hat{a}^H(\eta,\theta)\| &\le C_\ell P_A\left(\frac{|\theta|_*}{H}\right)s_A^{-\ell}(H/|e^{i\eta_j}-1|)^\ell(s_A/H)^d \\ &\le C_\ell H^{\ell-d}P_A\left(\frac{|\theta|_*}{H}\right)s_A^{d-\ell}|\eta|_*^{-\ell}.\end{aligned} \tag{9.24}$$

Or, by (9.19),

$$\|\hat{a}_0^H(\theta)^{-1}\hat{a}^H(\eta,\theta')\| \le C_\ell\left[1+\left(\frac{|\theta'|_*}{|\theta|_*}\right)^\kappa\right]H^{(1-\sigma_A-1|\eta|_*)-\ell} \tag{9.25}$$

where κ is independent of ℓ.

Having expressed the assumptions, we proceed to derive (9.6), but *first in a simplified form in which $(L_k^H)^{-1}$ replaces $(A^H)^{-1}$*. By (9.7) and (9.17),

$$\begin{aligned}(A^Hv^H)_\alpha &= \int\int_3 \hat{a}^H(\eta,\theta)\hat{v}^H(\theta)e^{i(\theta+\eta)\alpha}d\theta\,d\eta \\ &= \int\int_3 \hat{a}^H(\eta,\theta-\eta)\hat{v}^H(\theta-\eta)e^{i\theta\alpha}d\theta\,d\eta,\end{aligned} \tag{9.26}$$

where $\int_3$ denote integration over a domain $\pi \ge |\sigma|_* \ge CH^{1-\sigma_3}$, in which of course $|\theta|_* = |\theta|$. The latter equality in (9.26) is obtained by shifting θ by η (i.e., substituting θ for $\theta+\eta$). This shift leaves the integration domain effectively

unchanged since, by (9.24) and (9.11), $\hat{a}^H(\eta,\theta)$ practically vanishes (faster than any desired power of H) for $|\eta|_* \geq O(H^{1-\sigma_3})$. Hence, by (9.13) and (9.17),

$$(L_k^H - A^H)v_\alpha^H = \int\int_3 [\hat{a}^H(\eta,\theta)\hat{v}^H(\theta) - \hat{a}^H(\eta,\theta-\eta)\hat{v}^H(\theta-\eta)]e^{i\theta\alpha}d\theta\, d\eta,$$

and hence

$$(L_k^H)^{-1}(L_k^H - A^H)v_\alpha^H = T_\alpha^1 - T_\alpha^2 + T_\alpha^3 + T_\alpha^4, \tag{9.27}$$

where, for σ_B specified later,

$$\begin{aligned}
T_\alpha^1 &= \int_3 e^{i\theta\alpha}\hat{a}_0^H(\theta)^{-1}\int_{|\eta|\geq H^{1-\sigma_B}} \hat{a}^H(\eta,\theta)\hat{v}^H(\theta)\,d\eta d\theta \\
T_\alpha^2 &= \int_3 e^{\subset\theta\alpha}\hat{a}_0^H(\theta)^{-1}\int_{|\eta|\geq H^{1-\sigma_B}} \hat{a}^H(\eta,\theta-\eta)\hat{v}^H(\theta-\eta)d\eta\, d\theta \\
T_\alpha^3 &= \int_3 e^{i\theta\alpha}\hat{a}_0^H(\theta)^{-1}\int_{|\eta|\leq H^{1-\sigma_B}} \hat{a}^H(\eta,\theta-\eta)[\hat{v}^H(\theta) - \hat{v}^H(\theta-\eta)]d\eta\, d\theta \\
T_\alpha^4 &= \int_3 e^{i\theta\alpha}\hat{a}_0^H(\theta)^{-1}\int_{|\eta|\leq H^{1-\sigma_B}} [\hat{a}^H(\eta,\theta) - \hat{a}^H(\eta,\theta-\eta)]\hat{v}^H(\theta)d\eta\, d\theta.
\end{aligned}$$

In all these integrations $|\eta| \leq \pi$, hence $|\eta|_* = |\eta|$. By the Parseval identity, then the Cauchy-Schwarz inequality and then (9.25), for any $\ell > d/2$ and for some κ_1 independent of ℓ

$$\begin{aligned}
\|T^1\|^2 &= \int_3 \|\hat{a}_0^H(\theta)^{-1}\int_{|\eta|\geq H^{1-\sigma_B}} \hat{a}^H(\eta,\theta)\hat{v}^H(\theta)d\eta\|^2\, d\theta \\
&\leq C\|v^H\|^2 \int\int_{|\eta|\geq H^{1-\sigma_B}} \|\hat{a}_0^H(\theta)^{-1}\hat{a}^H(\eta,\theta)\|^2 d\eta\, d\theta \\
&\leq C_\ell \|v^H\|^2 H^{\kappa_1 + 2\ell(\sigma_B - \sigma_A)}.
\end{aligned}$$

Similarly

$$\begin{aligned}
\|T^2\| &\leq C\int\left(\int_{|\eta|\geq H^{1-\sigma_B}} \|\hat{a}_0^H(\theta)^{-1}\hat{a}^H(\eta,\theta)\|^2 d\eta\right)\int \|\hat{v}^H(\theta-\eta)\|^2 d\eta\, d\theta \\
&\leq C_\ell \|v^H\|^2 H^{\kappa_2 + 2\ell(\sigma_B - \sigma_A)}.
\end{aligned}$$

Again, by the Parseval identity, then the Cauchy-Schwarz inequality, then (9.22) (which is applicable here since we will choose $\sigma_B < \sigma_3$, hence $|\eta| \ll |\theta|$ in the

integrations below), and then (9.15).

$$\begin{aligned}
\|T^3\|^2 &= \int_3 \| \int_{|\eta|\le H^{1-\sigma_B}} \hat a_0^H(\theta)^{-1}\hat a^H(\eta,\theta-\eta)[\hat v^H(\theta)-\hat v^H(\theta-\eta)]d\eta\|^2 d\theta \\
&\le C\int_3\left(\int_{|\eta|\le H^{1-\sigma_B}} \|\hat a_0^H(\theta)^{-1}\hat a^H(\eta,\theta-\eta)\|^2 d\eta\right) \\
&\times\left(\int_{|\eta|\le H^{1-\sigma_B}} \|\hat v^H(\theta)-\hat v^H(\theta-\eta)\|^2 d\eta\right)d\theta \\
&\le CH^{2(\sigma_A-1)d}H^{d(1-\sigma_B)}\int_{|\eta|_*\le H^{1-\sigma_B}}\int \|\hat v^H(\theta)-\hat v^H(\theta-\eta)\|^2 d\theta\, d\eta \\
&\le CH^{d(2\sigma_A-\sigma_B-1)+2(\sigma_p-1)}\|v^H\|^2\int_{|\eta|\le H^{1-\sigma_B}}|\eta|^2\, d\eta \\
&\le CH^{2[d\sigma_A+\sigma_p-(d+1)\sigma_B]}\|v^H\|^2.
\end{aligned}$$

Finally, by the Parseval identity and (9.21)

$$\begin{aligned}
\|T^4\|^2 &= \int_3 \|\hat v^H(\theta)\|^2\int_{|\eta|\le H^{1-\sigma_B}}\hat a_0^H(\theta)^{-1}[\hat a^H(\eta,\theta)-\hat a^H(\eta,\theta-\eta)]\,d\eta\|^2 d\theta \\
&\le CH^{2[d(\sigma_A-1)+(d+1)(1-\sigma_B)-(1-\sigma_3)]}\|v^h\|^2 \\
&= CH^{2[d\sigma_A+\sigma_3-(d+1)\sigma_B]}\|v^H\|^2
\end{aligned}$$

By (9.11) we can choose σ_B such that $\sigma_A<\sigma_B<(d\sigma_A+\sigma_p)/(d+1)$, hence $(d+1)\sigma_B<d\sigma_A+\sigma_p<d\sigma_A+\sigma_3$ and $\sigma_B<\sigma_3$. Then, by choosing ℓ sufficiently large, the final exponents of H in the estimates for $\|T^1\|$, $\|T^2\|$, $\|T^3\|$ and $\|T^4\|$ are all positive. Hence, by (9.27), for some $\sigma_D>0$,

$$\|Bv^H\|\le CH^{\sigma_D}\|v^H\| \tag{9.28}$$

where

$$B=(L_k^H)^{-1}(L_k^H-A^H)=I-(L_k^H)^{-1}A^H. \tag{9.29}$$

Next, observe that the application of the operators A^H and $(L_k^H)^{-1}$, and hence also of B, leaves practically invariant the space of functions defined by (9.7)–(9.8). Indeed, by (9.26) and (9.24), for any $\sigma_3>\sigma_q>\sigma_A$, if v^H satisfies (9.7) then A^Hv^H (upto an error smaller than any desired power of h) has Fourier components only in the range $|\theta|\ge h^{1-\sigma_3}[1-O(h^{\sigma_3-\sigma_q})]$. Since A^H is local it surely preserves (9.8). The operator $(L_k^H)^{-1}$ certainly preserves (9.7), since it has constant coefficients. To show that it also preserves (9.8) one has to repeat the argument of Sec. (8.3), which indeed proved that, upon operating by $(L_k^H)^{-1}$, the support of any function is extended at most by $O(h^{\sigma_1})$, for any $\sigma_1<\sigma_3$. Using this for $\sigma_3>\sigma_1>\sigma_p$ it follows that the support (9.8) changes by as small a fraction as one wishes.

This implies that (9.28) can repeatedly be applied. More specifically, it implies that for any $m > 0$ and $\epsilon_* > 0$ one can choose H sufficiently small and N sufficiently large so that

$$\|B^n v^H\| \le (\epsilon_*/2)^n \|v^H\|, \qquad \text{for } \; n = 1, \dots, N-1$$

and so that

$$\|B^N v^H\| \le C H^m \|v^H\|.$$

Using these two inequalities, together with (6.13) and the identity

$$(A^H)^{-1}(L_k^H - A^H) = B + B^2 + \cdots + B^{N-1} + (A^H)^{-1} L_k^H B^N,$$

yields (9.6).

9.2.2. Proof of Claim III. Given v^H of the form (9.7) and $\varphi^{h;k}(x)$ satisfying (9.2), and the constant-coefficient elliptic operator $L = L_k^K$, we need to show, that there exists $\sigma > 0$ such that

$$\sum_k \|\varphi^{h;k} L^{-1} v^H - L^{-1}(\varphi^{h;k} v^H)\|^2 \le C H^\sigma \|L^{-1} v^H\|^2. \tag{9.30}$$

The ellipticity condition implies that the symbol $\hat{L}$ of L satisfies

$$\|[\hat{L}(\theta)^{-1} - \hat{L}(\theta+\eta)^{-1}]\hat{v}\| \le C \frac{|\eta|}{|\theta|} \|\hat{L}(\theta)^{-1}\hat{v}\| \tag{9.31}$$

for any $2|\eta| \le |\theta| \le \pi$ and any q-vector $\hat{v}$.

To prove (9.30) we use the Fourier expansion of $\varphi^{h;k}$ on grid h

$$\varphi_\alpha^k \underset{\text{def}}{=} \varphi^{h;k}(\alpha H) = \int_{|\eta| \le \pi} \hat{\varphi}^k(\eta) e^{i\eta\alpha} d\eta \tag{9.32}$$

where

$$\hat{\varphi}^k(\eta) = (2\pi)^{-d} \sum_\alpha \varphi_\alpha^{h;k} e^{-i\eta\alpha}.$$

This $\sum_\alpha$ includes a summation over α_{j_η}, where $|\eta_{j_\eta}| = |\eta|$. Summing that particular summation ℓ times by parts and using (9.2e) one obtains

$$|\hat{\varphi}^k(\eta)| \le C_\ell |\eta|^{-\ell} (H/s_p)^{\ell - d}, \tag{9.33}$$

for any positive integer ℓ. By (9.2a) we can choose σ_q and σ_r such that $\sigma_p < \sigma_q < \sigma_r < \sigma_3$. By (9.33),

$$|\hat{\varphi}^k(\eta)| \le C_\ell H^{\ell(\sigma_q - \sigma_p) + d(\sigma_p - 1)} \quad \text{for any} \quad |\eta| \ge H^{1-\sigma_q}$$

hence the contribution of such components is smaller than any power of H, and can thus be neglected. We can thus replace (9.32) by

$$\varphi_\alpha^k - \int_{|\eta|\le H^{1-\sigma_q}} \hat\varphi^k(\eta) e^{i\eta\alpha}\, d\eta. \tag{9.34}$$

By (9.7) we can partition v^H in the form

$$v_\alpha^H = \sum_j v_\alpha^{Hj},$$

with

$$v_\alpha^{Hj} = \int_{Dj} \hat v^H(\theta) e^{i\theta\alpha} d\theta,$$

where the domains D_j have the following properties:

(i) $D_j \subseteq \{\theta\colon |\theta - \theta^j| \le H^{1-\sigma_r}\}$, $\quad .5H^{1-\sigma_3} \le |\theta^j| \le \pi$.

(ii) The volume of D_j is $O(H^{d(1-\sigma_r)})$.

(iii) For the $j \ne k$, the volume of $D_j \cap D_k$ is at most $O(H^{(d-1)(1-\sigma_r)+1-\sigma_q})$.

By (9.34), the Fourier components of $L^{-1}\varphi^k v^{Hj}$, as well as those of $\varphi^k L^{-1} v^{Hj}$, can also be assumed to be contained in D_j (enlarging the latter at most by $O(H^{1-\sigma_r})$ in each direction, hence still observing the above properties). Since the volume of $D^\cap = \cup_{j\ne k}(D_j \cap D_k)$ is $O(H^{\sigma_r-\sigma_q})$, the domains D_j could in fact further be chosen so that the relative contribution of components from $D^\cap$ to the square norms on either side of (9.30) is only $O(H^{\sigma_r-\sigma_q})$, hence negligible. It is therefore enough to prove (9.30) for each v^{Hj} instead of v^H. Hence it is enough to prove (9.30) for v^H of the form $v_\alpha^H = w^H e^{i\theta\alpha}$, for some $|\theta| \ge H^{1-\sigma_3}$, where

$$w_\alpha^H = w^H(\alpha H) = \int_{|\eta|\le H^{1-\sigma_r}} \hat w(\eta) e^{i\eta\alpha} d\eta. \tag{9.35}$$

For such v^H we first have, by the Parseval identity and then by (9.31)

$$\begin{aligned}
\|\hat L(\theta)^{-1} v^H - L^{-1} v^H\|^2 &= \int_{|\eta|\le H^{1-\sigma_r}} [\hat L(\theta)^{-1} - \hat L(\theta+\eta)^{-1}] \hat w(\eta) e^{i(\theta+\eta)} d\eta\|^2 \\
&= \int_{|\eta|\le H^{1-\sigma_r}} \|[\hat L(\theta)^{-1} - L(\theta+\eta)^{-1}]\hat w(\eta)\|^2 d\eta \\
&\le C H^{2(\sigma_3-\sigma_r)} \|\hat L(\theta)^{-1} v^H\|^2.
\end{aligned} \tag{9.36}$$

Similarly (i.e., like (9.36), but with w^H being replaced by $\varphi^k w^H$)

$$\|\hat L(\theta)^{-1}\varphi^k v^H - L^{-1}(\varphi^k v^H)\|^2 \le C H^{2(\sigma_3-\sigma_r)} \|\hat L(\theta)^{-1}\varphi^k v^H\|^2. \tag{9.37}$$

By (9.2c) and (9.36)

$$\sum_k \|\hat{L}(\theta)^{-1}\varphi^k v^H - \varphi^k L^{-1} v^H\|^2 = \|\hat{L}(\theta)^{-1} v^H - L^{-1} v^H\| \leq CH^{2(\sigma_3-\sigma_r)} \|\hat{L}(\theta)^{-1} v^H\|^2. \tag{9.38}$$

By (9.37) and then (9.2c)

$$\sum_k \|\hat{L}(\theta)^{-1}\varphi^k v^H - L^{-1}(\varphi^k v^H)\|^2 \leq CH^{2(\sigma_3-\sigma_r)} \sum_k \|\hat{L}(\theta)^{-1}\varphi^k v^H\|^2 = CH^{2(\sigma_3-\sigma_r)} \|\hat{L}(\theta)^{-1} v^H\|^2. \tag{9.39}$$

Using (9.38) and (9.39), and then (9.36) again, one immediately obtains (9.30) with $\sigma = 2(\sigma_3 - \sigma_r)$.

9.3 Algebraic singularities

Observe that the only coefficient smoothness condition really needed in Sec. 9.2.1 is that, for any coefficient $a(x)$ of A, and for any sufficiently high ℓ and ℓ-order derivative ∂^ℓ, there holds

$$|\partial^\ell a(x)| \leq C_\ell h^{\kappa - \sigma_A \ell}, \qquad \text{for any} \quad x \in \Omega^{(h)}, \tag{9.40}$$

where κ is independent of ℓ. As in Sec. 9.1, the only requirement on this $\Omega^{(h)}$ is that the volume of $\Omega - \Omega^{(h)}$ tends to zero as $h \to 0$. In particular, choosing $\Omega^{(h)} = \Omega_\rho$ (cf. (7.1)), with $\rho = h^{\sigma_1}$ and $\sigma_A > \sigma_1 > 0$, it is easy to see that (9.40) is satisfied by any coefficient of the form $a(x) = b(x) r(x)^\alpha$, where $b(x)$ is sufficiently smooth, $r(x)$ is the distance of x from a fixed point (or line, etc.) on the boundary, and α is an arbitrary real power. Indeed, since $r \geq Ch^{\sigma_1}$ in Ω_ρ, for any $\ell > \alpha$ we have

$$|\partial^\ell a(x)| \leq C_\ell r^{\alpha-\ell} \leq C_\ell h^{\sigma_1(\alpha-\ell)} \leq C_\ell h^{\sigma_A(\alpha-\ell)}.$$

Thus, any type of algebraic singularity in coefficients is admissible. The only restriction to watch, of course, is that the CGA condition (Sec. 6.1) is still satisfied.

9.4 Finite elements and non-uniform discretization

Uniform finite elements give rise to difference equations directly amenable to local mode analyses. Those analyses can then be rigorously justified by the methods described above. Moreover, for reasons similar to those in Secs. 9.1, the rigorous analysis can be extended to *piecewise* uniform grids, with tending-to-zero-volume discontinuities. Most application of finite elements, especially those

which one would multigrid, or indeed exactly of this kind: the elements are uniform within each of the initially-given superelements, modified sometimes near boundaries and singularities.

Piecewise uniform finite-difference grids could of course similarly be analyzed. Moreover, within each piece the grids (or the elements) can smoothly vary (e.g., by some transformation): this is equivalent to (and can be viewed as) having a uniform grid with smoothly varying equations, hence it can be analyzed by the methods of Sec. 9.2–9.3.

Systems with arbitrarily ordered elements cannot be analyzed by the local *mode* analysis. (They are also inherently *much* less efficient then the partially ordered ones, especially with multigrid solvers). But we believe that their analysis should better still be a *local* analysis, like the present one. The traditional finite element multigrid analyses lose much insight by being *global*. They lose sight of the central fact that the potential multigrid efficiency is essentially independent of boundaries, singularities, etc.

For completely disordered grids, a *quantitative* two-level rigorous analysis, with realistic but not sharp constants, is given in [AMGT]. It cannot be localized, nor can it be extended to a multi-level theory, because it deals with systems that violate the CGA condition.

References

[ABDP] R.E. Alcouffe, A. Brandt, J.E. Dendy, Jr., and J.W. Painter: The multigrid methods for the diffusion equation with strongly discontinuous coefficients. *SIAM J. Sci. Stat. Comp.* **2** (1981), 430–454.

[AF85] A. Brandt: Multilevel adaptive techniques for compressible flow problems, Phase IV. (Annual Scientific Report, Air Force Grant AFOSR-84-0070), April, 1985.

[AMGT] A. Brandt: Algebraic multigrid theory: The symmetric case. *Appl. Math. Comp.* **19** (1986), 23–56.

[Bai] D. Bai and A. Brandt: Local mesh refinement multilevel techniques. *SIAM J. Sci. Stat. Comp.* **8** (1987), 109–134.

[BD] A. Brandt and N. Dinar: Multigrid solutions to elliptic flow problems. Numerical Methods for Partial Differential Equations (S. Parter, ed.), Academic Press 1979, pp. 53–147.

[Brs] D. Braess: The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation (revised). *Math. Comp.* **42** (1984), 368–386.

[CM3] A. Brandt: Multilevel computations: review and recent developments. Proc. 3rd Copper-Mountain Conf. on Multigrid Methods (S. McCormick,

ed), Marcel Dekker, 1988, pp. 220–312.

[G] Either [G82] or [G84].

[G82] A. Brandt: Guide to Multigrid Development. Multigrid Methods (Proc., Köln-Porz 1981. W. Hackbusch and U. Trottenberg, eds.) *Lecture Notes in Math.* **960**, Springer-Verlag.

[G84] A. Brandt: Multigrid Techniques: 1984 Guide, with Applications to Fluid Dynamics. Available as GMD Studien Nr. 85, GMD-AIW, Postfach 1240, D-5205, St. Augustin 1, W. Germany, 1984.

[H] W. Hackbusch: Multigrid Methods and Applications. Springer-Verlag, 1985.

[KW] A. Niestegge and K. Witsch: Analysis of a multigrid Stokes solver. to be submitted to AMC.

[MMB] J. Mandel, S. McCormick and R. Bank: Variational multigrid theory. Multigrid Methods (S. McCormick, ed.) Frontiers in Appl. Math., *SIAM*, Philadelphia, 1987, 131–178.

[MOC] A. Brandt: Multi-level adaptive solutions to boundary value problems. *Math. Comp.* **31** (1977), 333–390.

[ZN] U. Rüde and C. Zenger: On the treatment of singularities in the multigrid method. Multigrid Methods II (Proc., Cologne 1985. W. Hackbusch and U. Trottenberg, eds.), *Lecture Notes in Math.* **1228**, Springer-Verlag, Berlin 1986, pp. 262–272.