Chapter 6

Convergence proofs for multigrid methods

6.1 The structure of a TGM convergence proof (Hackbusch)

In order to construct a framework in which convergence of several multigrid algorithms can be proved we first give sufficient conditions for convergence of a two level algorithm (cf. [9]).

Definition 6.1 Let G_h be the approximate inverse related to some relaxation process for the solution of the equation $L_h u_h = f_h$. The relaxation process is said to have a *proper smoothing property of order* α if

$$\|L_h(I - G_h L_h)^\nu\| \le C_0(\nu) h^{-\alpha} \tag{6.1}$$

with $C_0(\nu)$ independent of h, $C_0(\nu) \to 0$ as $\nu \to \infty$. (This definition is a slightly simplified version of the one given in ([9])

Remark: Noticing that we have for the amplification operator of the error and residual of the relaxation respectively

$$M_h^{REL} = I_h \; - \; G_h L_h,$$

and

$$\overline{M}_h^{REL} = I_h \; - \; L_h G_h,$$

the property (6.1) can also be written as

$$\|L_h(M_h^{REL})^\nu\| \le C_0(\nu) h^{-\alpha} \tag{6.2}$$

or

$$\|(\overline{M}_h^{REL})^\nu L_h\| \le C_0(\nu) h^{-\alpha}. \tag{6.3}$$

Example 6.2 *The model problem*
In Chapter 4 we saw for the example operator L_h together with the damped Jacobi method as a relaxation method

$$\|L_h(I - G_h L_h)^\nu\| = \frac{4}{h^2} \sup_m \left| \begin{pmatrix} S_m^2 & 0 \\ 0 & C_m^2 \end{pmatrix} \begin{pmatrix} C_m^2 & 0 \\ 0 & S_m^2 \end{pmatrix}^\nu \right| =$$

$$= \frac{4}{h^2} \sup_m \left| \left(\begin{array}{cc} C_m^{2\nu} S_m^2 & 0 \\ 0 & S_m^{2\nu} C_m^2 \end{array} \right) \right| \le$$
$$\le \frac{4}{h^2} \max_{t\in[0,1]} t(1-t)^\nu \approx \frac{4}{h^2} \cdot \frac{1}{\nu e} \quad \text{for} \quad \nu \to \infty \, .$$

Hence, for this operator L_h, damped Jacobi has the proper smoothing property of order 2.

Theorem 6.3 In a two level algorithm, if

- (i) the operators L_h and L_H are relatively convergent of order α,
- (ii) the relaxation process for L_h satisfies a proper smoothing property of order α,
- (iii) the ratio H/h is bounded, and
- (iv) $\|M_h^{REL}\| \le C$ uniformly in h,

then, for the error-amplification operator of a two-level cycle with p pre- and q post-relaxation steps, $M_h^{TLA,p,q}$, we have the estimate

$$\|M_h^{TLA,p,q}\| \le C \cdot C_0(p),$$

where C is independent of h and $C_0(p) \to 0$ as $p \to \infty$.
Proof:

$$\begin{array}{rl} M_h^{TLA,p,q} & = (M_h^{REL})^q \, M_h^{CGC} (M_h^{REL})^p \\ & = (M_h^{REL})^q \, (I_h - P_{hH} L_H^{-1} \bar{R}_{Hh} L_h)(M_h^{REL})^p \\ & = (M_h^{REL})^q \, (L_h^{-1} - P_{hH} L_H^{-1} \bar{R}_{Hh}) L_h (M_h^{REL})^p \end{array}$$

so

$$\|M_h^{TLA,p,q}\| \quad \le \quad \|M_h^{REL}\|^q \|L_h^{-1} - P_{hH} L_H^{-1} \bar{R}_{Hh}\| \; \|L_h (M_h^{REL})^p\| \, .$$

$$\begin{array}{ll} (i) \implies & \|L_h^{-1} - P_{hH} L_H^{-1} R_{Hh}\| \le C \cdot H^\alpha \, , \\ (ii) \implies & \|L_h (M_h^{REL})^p\| \le C_0(p) \cdot h^{-\alpha} \, , \\ (iv) \implies & \|M_h^{REL}\|^q \le C. \end{array}$$

Thus, by (iii) it directly follows that

$$\|M_h^{TLA,p,q}\| \le C \cdot (H/h)^\alpha \cdot C_0(p) \le C \cdot C_0(p)$$

■

Theorem 6.4 If the conditions (i), (ii) and (iii) of Theorem 6.3 are satisfied, and instead of (iv) we have

$$\|\overline{M}_h^{REL}\| \le C \quad \text{uniformly for all} \quad h \, ,$$

then for the residual-amplification operator $\overline{M}_h^{TLA,p,q}$ of the two-level algorithm with p pre- and q post-relaxation steps, we have

$$\|\overline{M}_h^{TLA,p,q}\| \le C \, C_0(q),$$

where C is independent of h and $C_0(q) \to 0$ as $q \to \infty$.
Proof: Completely analogous to the proof of Theorem 6.3. ■

Remark: Notice that the estimate for $\|M_h^{TLA,p,q}\|$ does not depend on q whereas the estimate for $\|\overline{M}_h^{TLA,p,q}\|$ does not depend on p.

6.2 The structure of a MGM convergence proof (Hackbusch)

Whereas in the two-level algorithm we have to solve a coarse-grid problem in each iteration step, in the multi-level algorithm we solve this problem approximately, by the application of σ iteration steps of the same multi-level algorithm on the coarse level. In other words: a multi-level algorithm is a two-level algorithm in which the application of L_H^{-1} is replaced by σ multi-level algorithm sweeps on the H-level; i.e. L_H^{-1} is replaced by $[I-(M_H^{MGM})^\sigma]L_H^{-1}$, cf. Section 1.4. We will now give sufficient conditions for the convergence of a multi-level algorithm (cf. ([9]).

In the following we denote the error amplification operator of a multi-level iteration step on the h-level of discretisation by M_h^{MLA}, or - if we want to specify the strategy - we denote it by $M_h^{MLA,p,q,\sigma}$. The same operator on the next coarser level is denoted by M_H^{MLA}.

Theorem 6.5 Suppose the conditions (i), (ii) and (iii) of Theorem 6.3 are satisfied and in addition

- (v) P_{hH} is bounded and bounded from below, uniformly in h,
- (vi) $(M_h^{REL})^p$ and $(M_h^{REL})^q$ are bounded uniformly in h,
- (vii) p is sufficiently large,

then

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

Proof:

$$\begin{aligned}
&M_h^{MLA,p,q,\sigma}\\
&= (M_h^{REL})^q(I-P_{hH}(I-(M_H^{MLA})^\sigma)L_H^{-1}\bar R_{Hh}L_h)(M_h^{REL})^p\\
&= M_h^{TLA,p,q} + (M_h^{REL})^q P_{hH}(M_H^{MLA})^\sigma L_H^{-1}\bar R_{Hh}L_h(M_h^{REL})^p
\end{aligned}$$

so

$$\begin{aligned}
\|M_h^{MLA,p,q,\sigma}\| \le\ & \|M_h^{TLA,p,q}\| + \\
&+\|(M_h^{REL})^q\|\cdot\|P_{hH}\|\cdot\|M_H^{MLA}\|^\sigma\cdot\|L_H^{-1}\bar R_{Hh}L_h(M_h^{REL})^p\| .
\end{aligned} \tag{6.4}$$

Further

$$\begin{aligned}
-\ & P_{hH}L_H^{-1}\bar R_{Hh}L_h(M_h^{REL})^p\\
&= (L_h^{-1} - P_{hH}L_H^{-1}\bar R_{Hh} - L_h^{-1})L_h(M_h^{REL})^p\\
&= (L_h^{-1} - P_{hH}L_H^{-1}\bar R_{Hh})L_h(M_h^{REL})^p - (M_h^{REL})^p\\
&= M_h^{TLA,p,0} - (M_h^{REL})^p.
\end{aligned}$$

Since P_{hH} is bounded from below uniformly in h, according to (v), some $C>0$ exists, independent of h, such that

$$\|L_H^{-1}\bar R_{Hh}L_h(M_h^{REL})^p\| \le C\|M_h^{TLA,p,0} - (M_h^{REL})^p\|$$

Substituting this in (6.4) gives

$$\begin{aligned}
&\|M_h^{MLA,p,q,\sigma}\|\\
\le\ & \|M_h^{TLA,p,q}\| + C\|(M_h^{REL})^q\|\cdot\|P_{hH}\|\cdot\|M_H^{MLA}\|^\sigma\cdot\|M_h^{TLA,p,0} - (M_h^{REL})^p\|\\
\le\ & \|M_h^{TLA,p,q}\| + C\|(M_h^{REL})^q\|\cdot\|P_{hH}\|\cdot\|M_H^{MLA}\|^\sigma(\|M_h^{TLA,p,0}\| + \|(M_h^{REL})^p\|)
\end{aligned}$$

$$\begin{array}{lll}
\text{(i), (ii), (iii)} & \Longrightarrow & \|M_h^{TLA,p,0}\| \leq C \cdot C_0(p); \quad \text{cf. proof of Theorem6.3} \\
\text{(v)} & \Longrightarrow & \|P_{hH}\| \leq C, \\
\text{(vi)} & \Longrightarrow & \|(M_h^{REL})^q\| \leq C, \quad \|(M_h^{REL})^p\| \leq C \\
\text{(vii)} & \Longrightarrow & C_0(p) \leq C.
\end{array}$$

Now it directly follows that

$$\|M_h^{MLA,p,q,\sigma}\| \leq \|M_h^{TLA,p,q}\| + C\|M_H^{MLA}\|^\sigma.$$

■

We can obtain a similar result for the amplification operator of the residual:

Theorem 6.6 Suppose the conditions (i), (ii) and (iii) of Theorem 6.3 are satisfied and in addition

- (viii) $\bar{R}_{Hh}$ is bounded and bounded from below uniformly in h,
- (ix) $(\overline{M}_h^{REL})^p$ and $(\overline{M}_h^{REL})^q$ are bounded uniformly in h,
- (x) q is sufficiently large.

Then

$$\|\overline{M}_h^{MLA,p,q,\sigma}\| \leq \|\overline{M}_h^{TLA,p,q}\| + C\|\overline{M}_H^{MLA}\|^\sigma.$$

Proof: The proof is similar to the proof of Theorem 6.5. ■

From the Theorem 6.3 and 6.4 in combination with the hypotheses of Theorem 6.5 and Theorem 6.6 we may immediately conclude that

$$\|M_h^{TLA,p,q}\| \leq C \cdot C_0(p),$$

and

$$\|\overline{M}_h^{TLA,p,q}\| \leq C \cdot C_0(q).$$

We now denote the sequence of discretisations used in the multilevel algorithm by

$$L_{h_k} u_{h_k} = f_{h_k}, \quad k = 0, 1, 2, ...$$

with $|h_0| > |h_1| > |h_2| > ...$, or briefly by

$$L_k u_k = f_k, \quad k = 0, 1, 2, ...$$

Here k is the level of the discretisation. We define

$$\kappa_k = \|M_k^{TLA}\|, \quad k = 0, 1, 2, ... \tag{6.5}$$

and

$$\tilde{\kappa}_k = \|M_k^{MLA}\|, \quad k = 0, 1, 2, ... \tag{6.6}$$

If the conditions from Theorem 6.5 hold for all levels of the discretisation, then we can derive the following recursive relation for $\tilde{\kappa}_k$

$$\begin{cases} \tilde{\kappa}_1 = \kappa_1 \, , \\ \tilde{\kappa}_k \leq \kappa_k + C\tilde{\kappa}_{k-1}^\sigma , \end{cases} \tag{6.7}$$

where $\kappa_k \leq C \cdot C_0(p)$ independent of k. It can be proved that for $C_0(p)$ small enough and σ large enough (e.g. $\sigma = 2$), we have $\tilde{\kappa}_k \leq C < 1$. A similar recursion can be obtained for $\|\overline{M}_k^{MLA}\|$.

For a concrete multi-level algorithm used for solving e.g. a boundary value problem, convergence can be proved with the help of the above theorems. Here we will not go into the details of verifying the hypotheses of the theorems.

6.3 Tools for a MGM convergence proof

In the following sections we will give a MGM convergence proof following Bank and Douglas [2]. In this section we introduce several basic notions necessary for this proof. Let $\mathcal{H}$ be a Hilbert space, $a(\cdot,\cdot)$: $\mathcal{H}\times\mathcal{H}\to I\!R$ be a continuous, symmetric, positive definite, bilinear form, and $g(\cdot):\mathcal{H}\to I\!R$ be a continuous linear functional. The *continuous problem* to be solved is now defined as:

Find $u\in\mathcal{H}$ such that

$$a(u,v)=g(v) \text{ for all } v\in\mathcal{H}. \tag{6.8}$$

We will look for approximate solutions in nested finite dimensional spaces:

$$M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_j \subset \mathcal{H},$$

$$\dim\,(M_k)=N_k.$$

The *discrete problem* now is formulated as:

Find $z\in M_j$ such that

$$a(z,v)=g(v) \text{ for all } v\in M_j. \tag{6.9}$$

Let the set $\{\omega_k|k=1,2,\dots,N_j\}$ form a basis for M_j, then (6.9) can be formulated as: find z such that

$$a(z,\omega_k)=g(\omega_k) \quad \text{for} \quad k=1,2,\dots,N_j; \tag{6.10}$$

or equivalently

$$A\vec{z}=\vec{g} \tag{6.11}$$

where A is the $N_j\times N_j$ matrix with $A_{i,k}=a(\omega_i,\omega_k)$; $\vec{g}$ is the N_j-vector with $g_i=g(\omega_i)$, and $\vec{z}$ is the unknown N_j-vector with the coefficients of z. The problem (6.11) will be solved by a multigrid method.

Relaxation (smoothing)

We now consider the smoothing procedure (the relaxation). Let $b_j(\cdot,\cdot):M_j\times M_j\to I\!R$ be some approximation to $a(\cdot,\cdot)$, the so called *smoother* associated with M_j, such that b_j like a is a continuous, symmetric, positive definite bilinear form. The corresponding approximation for the matrix A is B with

$$B_{ik}=b_j(\omega_i,\omega_k).$$

We furthermore assume that

$$\sup_{v\in M_j}\frac{a(v,v)}{b_j(v,v)}\le 1\,, \tag{6.12}$$

or equivalently

$$\sup_{\vec{v}\in I\!R^{N_j}}\frac{\vec{v}^TA\vec{v}}{\vec{v}^TB\vec{v}}\le 1. \tag{6.13}$$

Example 6.7 *The model problem*
In our example from Chapter 4 we had

$$A=L_h=\frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

and this matrix was approximated by

$$B = \frac{1}{\omega} D_h = \frac{2}{\omega h^2} I,$$

for some damping factor ω. We already know that A has eigenvalues $\frac{4}{h^2} S_m^2$ and B has only one eigenvalue, namely $\frac{2}{\omega h^2}$. Therefore,

$$\sup_{\vec{v} \in \mathbb{R}^{N_j}} \frac{\vec{v}^T A v}{\vec{v}^T B v} \le \max_m 2\omega S_m^2 .$$

Thus, (6.12) is satisfied when $\omega \le 1/2$.

Definition 6.8 Let S be a positive definite symmetric matrix with eigenvalue decomposition $S = \Phi \Lambda \Phi^{-1}$; Φ is the matrix with orthonormal eigenvectors of S such that $\Phi \Phi^T = I$ and Λ is the diagonal matrix with the eigenvalues of S.

We define a function f on S as follows:

$$f(S) = \Phi f(\Lambda) \Phi^{-1} .$$

Let us now consider the generalised eigenvalue problem

$$A\Psi = B\Psi\Lambda, \tag{6.14}$$

with Λ (diagonal) and Ψ unknown, A and B as before. If we take Φ such that

$$\Psi = B^{-1/2}\Phi ,$$

then (6.14) can be written as the symmmetric positive definite eigenvalue problem

$$(B^{-1/2} A B^{-1/2})\Phi = \Phi\Lambda. \tag{6.15}$$

It is clear that there exists an orthonormal system $\{\phi_i\}$ of eigenvectors of $B^{-1/2}AB^{-1/2}$ and a set of eigenvalues $\{\lambda_i\}$ with $\lambda_i > 0$, and thus Ψ and Λ can be determined. We have the following useful relations:

Lemma 6.9

$$\Psi^T A \Psi = \Lambda , \tag{6.16}$$

$$\Psi B \Psi = I , \tag{6.17}$$

$$0 < \lambda_1 \le \lambda_2 \le ... \le \lambda_{N_j} \le 1 . \tag{6.18}$$

Proof:

$$\begin{aligned} a) \quad \Psi^T A \Psi &= (B^{-1/2}\Phi)^T A (B^{-1/2}\Phi) \\ &= \Phi^T B^{-1/2} A B^{-1/2} \Phi \\ &= \Phi^T \Phi \Lambda \\ &= \Lambda . \end{aligned}$$

$$\begin{aligned} b) \quad \Psi^T B \Psi &= (B^{-1/2}\Phi)^T B (B^{-1/2}\Phi) \\ &= \Phi^T B^{-1/2} B B^{-1/2} \Phi) \\ &= \Phi^T \Phi \\ &= I . \end{aligned}$$

c) Let ψ_i be the i-th column of Ψ. hen, because of (6.16) and the fact that A is positive definite, we have

$$0 \, < \, \lambda_i = \psi_i^T A \psi_i \, .$$

Using (6.13) and (6.17) we obtain

$$0 \, < \, \lambda_i = \psi_i^T A \psi_i \, \leq \, \psi_i^T B \psi_i = 1.$$

Arranging the λ_i's gives

$$0 \, < \, \lambda_1 \, \leq \, \lambda_2 \, \leq .... \leq \, \lambda_{N_j} \, \leq 1.$$

■

Definition 6.10 On M_j we define a norm $\| \cdot \|_\theta$, $\theta \in I\!R$, by

$$\begin{aligned} \|v\|_\theta^2 \; &= \vec{v}^T B^{1/2} (B^{-1/2} A B^{-1/2})^\theta B^{1/2} \vec{v} \\ &= \vec{v}^T B^{1/2} \Phi \Lambda^\theta \Phi^{-1} B^{1/2} \vec{v} \\ &= \vec{c}^T \Lambda^\theta \vec{c} \end{aligned} \tag{6.19}$$

with $\vec{v} = \Psi \vec{c}$ and $(\vec{v})_i$ such that $v = \sum_{i=1}^{N_j} v_i \omega_i$.

Remark: It can be easily verified (use $\lambda_i > 0$), that the above definition really defines a norm.

Example 6.11 *The model problem*
We analyse the vector $\vec{c}$ in case of the example from Chapter 4. Again we take $\omega = 1/2$.

$$B = \frac{2}{\omega h^2} I = \frac{4}{h^2} I \quad \text{and hence} \quad B^{1/2} = \frac{2}{h} I \, ,$$

$$\vec{c} = \Psi^{-1} \vec{v} = \Phi^T B^{1/2} \vec{v} = \frac{2}{h} \Phi^T \vec{v}. \tag{6.20}$$

If we denote the i-th column of Φ by ϕ_i then we find for the i-th component of $\vec{c}$:

$$c_i = \frac{2}{h} < \phi_i, \vec{v} > .$$

Thus, in this case the i-th component of the vector c is a measure of the size of the component in the direction ϕ_i of $\vec{v}$.

As a consequence of the preceding definition we have the following

Lemma 6.12

$$\|v\|_0^2 = \vec{v}^T B \vec{v} \tag{6.21}$$

$$\|v\|_1^2 = \vec{v}^T A \vec{v} \tag{6.22}$$

$$|a(v,w)| \, \leq \, \|v\|_{1-\alpha} \|w\|_{1+\alpha}, \qquad \alpha \in I\!R. \tag{6.23}$$

Proof: (6.21) and (6.22) are trivial consequences of Definition 6.10. We will therefore prove only (6.23). If we write $v = \Sigma_i v_i \omega_i$, $w = \Sigma_i w_i \omega_i$ and denote the vectors with the coefficients v_i and w_i as usual by $\vec{v}$ respectively $\vec{w}$, we find

$$\begin{aligned} a(v,w) \; &= \vec{v}^T A \vec{w} \\ &= \vec{v}^T B^{1/2} \Phi \Lambda \Phi^{-1} B^{1/2} \vec{w}, \end{aligned}$$

the last equality following from (6.15).

If we write

$$\vec{c} = \Phi^{-1} B^{1/2} \vec{v} \quad (\text{or} \;\; \vec{v} = \Psi \vec{c}\,),$$
$$\vec{d} = \Phi^{-1} B^{1/2} \vec{w} \quad (\text{or} \;\; \vec{w} = \Psi \vec{d}\,),$$

we find

$$a(v,w) = \vec{c}^T \Lambda \vec{d} = \sum_i c_i \lambda_i d_i = \sum_i (c_i \lambda_i^{(1-\alpha)/2})\,(d_i \lambda_i^{(1+\alpha)/2}),$$

and with the use of the Cauchy-Schwartz inequality:

$$\begin{aligned} |a(v,w)| \quad &\le \sqrt{\textstyle\sum_i (c_i \lambda_i^{(1-\alpha)/2})^2}\;\sqrt{\textstyle\sum_i (d_i \lambda_i^{(1+\alpha)/2})^2} \\ &= \sqrt{\textstyle\sum_i c_i \lambda_i^{1-\alpha} c_i}\;\sqrt{\textstyle\sum_i d_i \lambda_i^{1+\alpha} d_i} \\ &= \sqrt{\vec{c}\Lambda^{1-\alpha}\vec{c}}\;\sqrt{\vec{d}^T \Lambda^{1+\alpha}\vec{d}} \\ &= \|v\|_{1-\alpha}\;\|w\|_{1+\alpha}. \end{aligned}$$

■

6.4 The effect of the smoothing operator

When solving problem (6.11) by a multigrid method, a smoothing operator will be applied several times. An approximate inverse B^{-1} of A is at our disposal. The smoothing operator is therefore defined as

$$S = I - B^{-1}A = B^{-1}(B - A). \tag{6.24}$$

We now want to analyse what happens to an error $e \in M_j$ when S is applied.

Theorem 6.13 Let $e \in M_j$, $\vec{c}^e$ such that $\vec{e} = \Psi \vec{c}^{\vec{e}}$, (cf. equation 6.20) or equivalently: $\vec{c}^{\vec{e}}$ such that $e = \sum_i c_i^e \, \psi_i$. With S as in (6.24) and Λ as in the preceding section we have

$$Se = \sum_i c_i^e (1 - \lambda_i)\, \psi_i. \tag{6.25}$$

Proof: Let $\vec{c}^S e$ be such that $S\vec{e} = \Psi \vec{c}^{Se}$, then

$$\begin{aligned} \vec{c}^{Se} \quad &= \Psi^{-1} S \vec{e} \\ &= \Phi^{-1} B^{1/2} S \vec{e} \\ &= \Phi^{-1} B^{1/2} B^{-1} (B - A)\vec{e} \\ &= \Phi^{-1} B^{1/2} \vec{e} - \Phi^{-1} B^{-1/2} A \vec{e} \\ &= \Psi^{-1} \vec{e} - \Phi^{-1} B^{-1/2} A \Psi \vec{c}^e \\ &= \vec{c}^e - \Phi^{-1} B^{-1/2} A B^{-1/2} \Phi \vec{c}^{\vec{e}} \\ &= [I - \Phi^{-1} B^{-1/2} A B^{-1/2} \Phi] \vec{c}^{\vec{e}} \\ &= [I - \Lambda] \vec{c}^e\,. \end{aligned}$$

So $c_i^{Se} = (1 - \lambda_i) c_i^e$ and hence

$$Se = \sum_i c_i^{Se}\, \psi_i = \sum_i c_i^e (1 - \lambda_i)\, \psi_i\,.$$

■

As in Chapter 4, we see that components having high frequencies (large λ_i) are damped very well, whereas components having low frequencies (small λ_i) remain almost unchanged.

As a consequence of Theorem 6.13 we find

Theorem 6.14 With f as in lemma 4.1, $\alpha \in \mathbb{R}$, p an integer we have

$$\|S^p e\|_\theta \leq f(p, \alpha/2) \;\; \|e\|_{\theta-\alpha}. \tag{6.26}$$

Proof: It follows from Theorem 6.13 that

$$S^p e = \sum_i c_i^e \, (1-\lambda_i)^p \, \psi_i \, ,$$

and thus

$$\vec{c}^{S^p e} = (I - \Lambda)^p \, \vec{c}^e \, .$$

By definition

$$\begin{aligned} \|S^p e\|_\theta^2 &= \vec{c}^{S^p e} \Lambda^\theta (\vec{c}^{S^p e})^T \\ &= \textstyle\sum_i (c_i^e)^2 (1-\lambda_i)^{2p} \lambda_i^\theta \\ &= \textstyle\sum_i (c_i^e)^2 \lambda_i^{\theta-\alpha} (1-\lambda_i)^{2p} \lambda_i^\alpha . \end{aligned}$$

By (6.18) $0 < \lambda_i \leq 1$ and therefore

$$\|S^p e\|_\theta^2 \leq \max_{0 \leq \lambda \leq 1} [(1-\lambda)^{2p} \lambda^\alpha] \sum_i (c_i^e)^2 \lambda_i^{\theta-\alpha} = f(2p, \alpha) \; \| e\|_{\theta-\alpha}^2 \, ,$$

or

$$\|S^p e\|_\theta \leq f(p, \alpha/2) \; \|e\|_{\theta-\alpha}.$$

■

Example 6.15 *The model problem*

We now want to apply the above theory to our example from Chapter 4. Again we take $\omega = 1/2$.

$$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}, \quad B = \frac{4}{h^2} I$$

and hence

$$B^{-1/2} A B^{-1/2} = \frac{1}{4} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & \\ & & -1 & 2 \end{pmatrix}$$

with eigenvalues $\lambda_m = S_m^2$ and corresponding eigenvectors

$$\phi_m = \begin{pmatrix} \sin(m\pi h) \\ \sin(2m\pi h) \\ \dots \\ \sin(N_j m \pi h) \end{pmatrix} ,$$

$$\Phi = (\phi_1, \phi_2, ..., \phi_{N_j}).$$

We have $\Phi = B^{1/2}\Psi$ and hence the following relation for the columns of Φ and Ψ is valid:

$$\phi_m = \frac{2}{h}\psi_m.$$

Now we are able to compute $\|v\|_\theta$ and $\|S^p v\|_\theta$ for $v \in M_j$. Let $v = \Sigma_m \alpha_m \phi_m \in M_j$, then

$$v = \sum_m \alpha_m \frac{2}{h}\psi_m\,,$$

and therefore (cf. Definition 6.10)

$$\begin{aligned}\|v\|_\theta^2 &= \textstyle\sum_m (\alpha_m \frac{2}{h})^2\,\lambda_m^\theta \\ &= \textstyle\sum_m \alpha_m^2\,\frac{4}{h^2}\,S_m^{2\theta}. \\ \|v\|_\theta &= \textstyle\frac{2}{h}\sqrt{\sum_m \alpha_m^2 S_m^{2\theta}}.\end{aligned}$$

In Chapter 4 we already observed that $S\phi_m = C_m^2\phi_m$ and hence

$$\begin{aligned}S^p v = \;& S^p \textstyle\sum_m \alpha_m\phi_m \\ &= \textstyle\sum_m \alpha_m S^p\phi_m \\ &= \textstyle\sum_m \alpha_m C_m^{2p}\phi_m \\ &= \textstyle\sum_m \alpha_m C_m^{2p}\cdot\frac{2}{h}\psi_m,\end{aligned}$$

thus

$$\begin{aligned}\|S^p v\|_\theta^2 = \;& \textstyle\sum_m (\alpha_m C_m^{2p}\cdot\frac{2}{h})^2\lambda_m^\theta \\ &= \textstyle\sum_m \alpha_m^2 C_m^{4p}\frac{4}{h^2}S_m^{2\theta}\,, \\ \|S^p v\|_\theta &= \textstyle\frac{2}{h}\sqrt{\sum_m \alpha_m^2 C_m^{4p} S_m^{2\theta}}.\end{aligned}$$

6.5 The effect of the coarse grid correction

We will denote the amplification operator of the error associated with the coarse grid correction by

$$C:\ M_j \;\rightarrow\; M_j.$$

In a coarse grid correction, a smooth error $v^{\rm old} \in M_j$ is approximated in M_{j-1}. This approximation is called v^{corr} and hence

$$v^{\rm new}\; - v^{\rm old}\; = -v^{corr} \in M_{j-1}$$

or

$$Cv^{\rm old}\; - v^{\rm old}\; \in M_{j-1}.$$

Because we consider a nested sequence of Galerkin discretisations (6.9), the result of the coarse grid correction is that the new error has no component left in M_{j-1}. Thus we have the following two properties of the coarse grid operator $C: M_j \;\rightarrow\; M_j$

$$Cv - v \in M_{j-1} \;\text{ for all }\; v \in M_j. \tag{6.27}$$

$$a(Cv, w) = 0 \;\text{ for all }\; w \in M_{j-1},\; v \in M_j. \tag{6.28}$$

Remark: Property (6.28) is also written as $Cv \in M_{j-1}^\perp$.

Lemma 6.16 If C satisfies (6.27) and (6.28) then

$$C^2 = C.$$

Proof: If $v \in M_j$ then $Cv \in M_j$, and hence by (6.27)

$$CCv - Cv \in M_{j-1}.$$

Using (6.28) with $v = Cv - v$, $w = CCv - Cv$ we find

$$a(CCv - Cv, CCv - Cv) = 0,$$

or by (6.22)

$$\|CCv - Cv\|_{\theta=1} = 0.$$

Thus $CCv = Cv$ for all $v \in M_j$. ∎

From now on we will use the following notation

$$|||\cdot||| = \|\cdot\|_{\theta=1}\,.$$

Lemma 6.17 Let C satisfy (6.27) and (6.28). If $\tilde{C}$ is such that

$$\tilde{C}v - v \in M_{j-1} \quad \text{for all} \quad v \in M_j,$$

and

$$|||\tilde{C}v - Cv||| \,<\, \gamma|||Cv - v||| \text{ for some}\, \gamma \,>\, 0,$$

then

$$|||\tilde{C}v \,+\, (\gamma - 1)Cv||| \,\leq\, \gamma|||v|||$$

Proof: For all $v \in M_j$

$$\begin{aligned}|||\tilde{C}v \,+\, (\gamma - 1)Cv|||^2 &= \\ &= a(\tilde{C}v + (\gamma - 1)Cv\,,\, \tilde{C}v + (\gamma - 1)Cv) = \\ &= a(\tilde{C}v - Cv\,,\, \tilde{C}v - Cv) + 2\gamma\, a(\tilde{C}v - Cv, Cv) + \gamma^2 a(Cv, Cv) = \\ &= |||\tilde{C}v - Cv|||^2 \,+\, \gamma^2|||Cv|||^2.\end{aligned}$$

The last equality being valid because

$$\tilde{C}v - Cv = (\tilde{C}v - v) - (Cv - v) \in M_{j-1}$$

and hence

$$a(\tilde{C}v - Cv, Cv) = 0.$$

Now the assumption that

$$|||\tilde{C}v - Cv||| \,<\, \gamma|||Cv - v|||$$

yields

$$|||\tilde{C}v + (\gamma - 1)Cv|||^2 \,\leq\, \gamma^2(|||Cv - v|||^2 + |||Cv|||^2) \,=\, \gamma^2|||v|||^2.$$

(Note that $a(Cv - v, Cv) = 0$ and thus we may use Pythagoras' rule.) Taking the square root on both sides we obtain the desired result. ∎

From now on we will assume that the following hypothesis is satisfied:

Hypothesis 6.18 For given $\alpha > 0$, and $\theta \in \mathbb{R}$, there exists some constant $\kappa \geq 1$ such that for all $v \in M_j \bigcap M_{j-1}^{\perp}$

$$\|v\|_{\theta-\alpha}^2 \leq \kappa^{\alpha} \|v\|_{\theta}^2 \quad \text{uniform for all } j. \tag{6.29}$$

Remark: If κ can be chosen such that $\frac{1}{\kappa} \leq \lambda_i \leq 1$ for all i, uniform in j, then hypothesis 6.18 holds for this κ. For if we take $\vec{v} = \Psi\vec{c}$, then we have the following equivalences

$$\|v\|_{\theta-\alpha} \leq \kappa^{\alpha/2} \|v\|_{\theta} \iff$$

$$\|v\|_{\theta-\alpha}^2 \leq \kappa^{\alpha} \|v\|_{\theta}^2 \iff$$

$$\sum_i c_i^2 \lambda_i^{\theta-\alpha} \leq \kappa^{\alpha} \sum_i c_i^2 \lambda_i^{\theta}.$$

Hence, if $\frac{1}{\kappa} \leq \lambda_i \leq 1$ then $\kappa^{\alpha} \geq \lambda_i^{-\alpha}$ and (6.29) is valid. Notice that in practice there is often no uniform positive lower bound for λ_i because λ_i depends on j.

If for given θ and α, a κ exists such that (6.29) holds, then

$$\kappa \geq \sup_j \sup_{v \in M_j \cap M_{j-1}^{\perp}} \frac{\|v\|_{\theta-\alpha}^{2/\alpha}}{\|v\|_{\theta}^{2/\alpha}} \tag{6.30}$$

and vice versa; if for given θ and α the supremum (6.30) exists, then (6.29) holds with κ as in (6.30).

Lemma 6.19 If hypothesis 6.18 is true for α_0 and θ_0 then it is true for α and θ_0 with $0 < \alpha \leq \alpha_0$.
Proof: This proof is left as an exercise. ∎

Example 6.20 *The model problem*
For our example in Chapter 4 we will determine some κ for $\theta = \alpha = 1$. To do this we have to know first which functions v are in $M_j \cap M_{j-1}^{\perp}$. In Chapter 4 we saw that for the amplification operator of the error, $C = M_h^{CGC*}$ corresponding to the coarse grid correction with $\bar{R}_{Hh}^*$ instead of $\bar{R}_{Hh}$

$$C(\phi_m) = S_m^2 \phi_m + S_m^2 \phi_{N-m},$$

$$C(\phi_{N-m}) = C_m^2 \phi_m + C_m^2 \phi_{N-m},$$

hence

$$C(\phi_m + \phi_{N-m}) = \phi_m + \phi_{N-m},$$

$$C(C_m^2 \phi_m - S_m^2 \phi_{N-m}) = 0.$$

Using (6.27) and (6.28) we find

$$M_{j-1} \supseteq \text{Span}\{C_m^2 \phi_m - S_m^2 \phi_{N-m}\}$$

and

$$M_{j-1}^{\perp} \supseteq \text{Span}\{\phi_m + \phi_{N-m}\}.$$

Counting the dimensions of the spaces it follows that

$$M_{j-1} = \text{Span}\{C_m^2 \phi_m - S_m^2 \phi_{N-m}\}$$

and

$$M_{j-1}^{\perp} = \text{Span}\{\phi_m + \phi_{N-m}\}.$$

Now

$$\kappa = \sup_{v \in M_j \cap M_{j-1}^\perp} \frac{\|v\|_{\theta-\alpha}^{2/\alpha}}{\|v\|_\theta^{2/\alpha}} = \sup_{v \in M_{j-1}^\perp} \frac{\|v\|_0^2}{\|v\|_1^2}\,.$$

Let $v = \sum_m \alpha_m \phi_m$, then $\|v\|_\theta^2 = \frac{4}{h}\sum_m \alpha_m^2 S_m^{2\theta}$ (cf. example 6.15) and hence

$$\kappa = \sup_{v \in M_{j-1}^\perp} \frac{\frac{4}{h}\sum_m \alpha_m^2}{\frac{4}{h}\sum_m \alpha_m^2 S_m^2} = \sup_{v \in M_{j-1}^\perp} \frac{\sum_m \alpha_m^2}{\sum_m \alpha_m^2 S_m^2} =$$

$$= \sup_{v \in M_{j-1}^\perp} \frac{\sum_{m=1}^{N/2}(\alpha_m^2 S_m^2 + \alpha_{N-m}^2)}{\sum_{m=1}^{N/2}(\alpha_m^2 S_m^2 + \alpha_{N-m}^2 S_{N-m}^2)}\,,$$

the last equality being valid because N is even (assumption from Chapter 4).

Since $M_{j-1}^\perp = \text{Span}\,\{\phi_m + \phi_{N-m}\}$ we have $\alpha_m = \alpha_{N-m}$ for $v = \Sigma_m \alpha_m \phi_m \in M_{j-1}^\perp$. Hence

$$\kappa = \sup_{(\alpha_1,\ldots,\alpha_{N/2})} \frac{\sum_{m=1}^{N/2}(\alpha_m^2 + \alpha_m^2)}{\sum_{m=1}^{N/2}(\alpha_m^2 S_m^2 + \alpha_m^2 C_m^2)} = \sup_{(\alpha_1,\ldots,\alpha_{N/2})} \frac{\sum_{m=1}^{N/2} 2\alpha_m^2}{\sum_{m=1}^{N/2} \alpha_m^2} = 2\,.$$

Lemma 6.21 The validity of our hypothesis has the following consequences

$$\|Cv\|_1 \;\le\; \kappa^{\alpha/2}\|v\|_{1+\alpha} \tag{6.31}$$

$$\|CS^p v\|_1 \;\le\; \kappa^{\alpha/2} f(p,\alpha/2)\,\|v\|_1 \tag{6.32}$$

$$\|S^q Cv\|_1 \;\le\; \kappa^{\alpha/2} f(q,\alpha/2)\,\|Cv\|_1 \tag{6.33}$$

$$\|S^q Cv\|_{1+\alpha} \;\le\; \kappa^{\alpha/2} f(q,\alpha/2)\,\|v\|_{1+\alpha} \tag{6.34}$$

Proof:
(6.31): $\|Cv\|_1^2 = a(Cv,Cv) = a(Cv,Cv-v+v) = a(Cv,Cv-v)+a(Cv,v) = a(Cv,v)$. The last equality follows from (6.27) and (6.28). Using (6.23) and (6.29) ($\theta = 1$) yields

$$a(Cv,v) \;\le\; \|Cv\|_{1-\alpha}\|v\|_{1+\alpha} \;\le\; \kappa^{\alpha/2}\|Cv\|_1\|v\|_{1+\alpha}\,,$$

or summarising

$$\|Cv\|_1^2 \;\le\; \kappa^{\alpha/2}\|Cv\|_1\|v\|_{1+\alpha}.$$

Division by $\|Cv\|_1$ gives the desired result.
(6.32): From (6.31) follows

$$\|CS^p v\|_1 \;\le\; \kappa^{\alpha/2}\|S^p v\|_{1+\alpha}\,.$$

Now Theorem 6.14 gives

$$\|CS^p v\|_1 \;\le\; \kappa^{\alpha/2} f(p,\alpha/2)\|v\|_1.$$

(6.33): The successive use of Theorem 6.14 and our hypothesis with $\theta = 1$ yields

$$\|S^q Cv\|_1 \;\le\; f(q,\alpha/2)\|Cv\|_{1-\alpha} \;\le\; k^{\alpha/2} f(q,\alpha/2)\|Cv\|_1.$$

(6.34): Now we use Theorem 6.14 and (6.31)

$$\|S^q Cv\|_{1+\alpha} \;\le\; f(q,\alpha/2)\|Cv\|_1 \;\le\; \kappa^{\alpha/2} f(q,\alpha/2)\|v\|_{1+\alpha}.$$

■

6.6 The convergence proof for the TGM

In a two-level algorithm the error-amplification operator is $S^q C S^p$. For this operator we have the following

Theorem 6.22 If the hypothesis (6.29) is true then

$$|||S^q C S^p v||| \leq \gamma |||v||| \text{ for all } v \in M_j, \tag{6.35}$$

where

$$\gamma = \inf_\alpha f(q, \alpha/2)\, f(p, \alpha/2)\, \kappa^\alpha,$$

the infimum taken over those α for which (6.29) is true.
Proof: The use of (6.33) and (6.34) results in

$$\begin{aligned} |||S^q C S^p v||| &\leq \kappa^{\alpha/2} f(q,\alpha/2)\, |||C S^p v||| \\ &\leq \kappa^{\alpha/2} f(q,\alpha/2) \kappa^{\alpha/2} f(p,\alpha/2)\, |||v||| \\ &\leq \gamma |||v||| \,. \end{aligned}$$

■

In the remainder of this chapter we restrict ourselves to $p = q$ and we write $m = p + q$. As a result of this restriction we have for γ as in Theorem 6.22:

$$\gamma = \inf_\alpha \,[f(p,\alpha/2)]^2\, \kappa^\alpha = \inf_\alpha \, f(2p,\alpha)\, \kappa^\alpha = \inf_\alpha \, f(m,\alpha)\, \kappa^\alpha \cdot$$

Lemma 6.23 If the hypothesis is true for κ such that $\alpha = m/(\kappa - 1)$, then the convergence factor γ of the two level algorithm equals $((\kappa-1)/\kappa)^m$.
Proof: We compute $\gamma = \inf_\alpha \, f(m,\alpha)\, \kappa^\alpha$.

$$f(m,\alpha)\kappa^\alpha = \frac{\kappa^\alpha m^m \alpha^\alpha}{(m+\alpha)^{m+\alpha}}$$

$$\inf_\alpha \kappa^\alpha f(m,\alpha) = \exp\left(\inf_\alpha \ln \frac{\kappa^\alpha m^m \alpha^\alpha}{(m+\alpha)^{m+\alpha}} \right) =$$

$$\exp(\inf_\alpha \,(\alpha \, \ln \kappa + m\, \ln m + \alpha \, \ln \alpha - (m+\alpha)\, \ln(m+\alpha)))\,.$$

Now the infimum will be determined by computing the zero of the derivative of

$$g(\alpha) = \alpha \ln \kappa + m \ln m + \alpha \ln \alpha - (m+\alpha)\ln(m+\alpha).$$

$$g'(\alpha) = \ln \kappa + \ln \alpha + 1 - \ln(m+\alpha) - 1 = \ln \frac{\kappa\alpha}{m+\alpha} \cdot$$

$$g'(\alpha) = 0 \iff \frac{\kappa\alpha}{m+\alpha} = 1 \iff \alpha = \frac{m}{\kappa-1} \cdot$$

$$g''(\alpha) = \frac{1}{\alpha} - \frac{1}{m+\alpha} = \frac{m}{\alpha(\alpha+m)} \iff g''(\alpha) > 0 \;\text{ for }\; \alpha > 0\,.$$

Hence, g takes its infimum at $\alpha = m/(\kappa-1)$.

We assumed the hypothesis being true for k such that $\alpha = m/(k-1)$. Since the infimum has to be taken over those α for which (6.29) holds, we have

$$\gamma \;=\; \inf_\alpha \, f(m,\alpha)\kappa^\alpha = \frac{\left(\frac{\kappa^m}{\kappa-1}\right)^{\frac{m}{\kappa-1}} m^m}{\left(m + \frac{m}{\kappa-1}\right)^m + \frac{m}{\kappa-1}} = \left(\frac{\kappa-1}{\kappa}\right)^m .$$

■

6.7 The convergence proof for the W-cycle and the V-cycle

In the following theorem we introduce the operator $\tilde{C}$, an approximation of the coarse grid correction operator C. The condition on $\tilde{C}$ is such that $\tilde{C}$ may represent a W-cycle approximation (or a V-cycle approximation) to the coarse grid correction.

Theorem 6.24 Suppose that the hypothesis (6.29) is true and that γ in (6.35) satisfies $\gamma < \frac{1}{2}$. Let $\tilde{C}$ be such that

$$|||\tilde{C}v - Cv||| \leq \tilde{\gamma}^2|||Cv - v||| \text{ for all } v \in M_j,$$

where

$$\frac{\gamma}{1-\gamma} \leq \tilde{\gamma} < 1.$$

Then

$$|||S^{m/2}\tilde{C}S^{m/2}v||| \leq \tilde{\gamma}|||v|||.$$

Proof:

$$\begin{aligned} |||S^{m/2}\tilde{C}S^{m/2}v||| &= |||S^{m/2}(1-\tilde{\gamma}^2)CS^{m/2}v + S^{m/2}(\tilde{C} + (\tilde{\gamma}^2 - 1)C)S^{m/2}v||| \\ &\leq (1-\tilde{\gamma}^2)\,|||S^{m/2}CS^{m/2}v||| + |||S^{m/2}(\tilde{C} + (\tilde{\gamma}^2-1)C)S^{m/2}v|||\,. \end{aligned} \tag{6.36}$$

Using Theorem 6.22 we have

$$|||S^{m/2}CS^{m/2}v||| \leq \gamma|||v|||\,.$$

Furthermore if we take $\alpha = 0$, $\theta = 1$ in Theorem 6.14 it follows that

$$|||S^{m/2}(\tilde{C} - (\tilde{\gamma}^2 - 1)C)S^{m/2}v||| \leq |||(\tilde{C} - (\tilde{\gamma}^2 - 1)C)S^{m/2}v|||\,.$$

In addition, application of Lemma 6.17 yields

$$|||(\tilde{C} - (\tilde{\gamma}^2 - 1)C)S^{m/2}v||| \leq \tilde{\gamma}^2|||S^{m/2}v|||\,,$$

and with Theorem 6.14 we find

$$\tilde{\gamma}^2|||S^{m/2}v||| \leq \tilde{\alpha}^2|||v|||\,.$$

Substitution of all this in (6.36) gives

$$|||S^{m/2}\tilde{C}S^{m/2}v||| \leq [(1-\tilde{\gamma}^2)\gamma + \tilde{\gamma}^2]\;|||v|||\,.$$

We assumed

$$\begin{gathered} \frac{\gamma}{1-\gamma} \leq \tilde{\gamma} < 1 \quad \Longleftrightarrow \\ \gamma \leq \tilde{\gamma} - \gamma\tilde{\gamma} \quad \Longleftrightarrow \\ \gamma(1+\tilde{\gamma})(1-\tilde{\gamma}) \leq \tilde{\gamma}(1-\tilde{\gamma}) \quad \Longleftrightarrow \\ \gamma - \gamma\tilde{\gamma}^2 \leq \tilde{\gamma} - \tilde{\gamma}^2 \quad \Longleftrightarrow \\ \tilde{\gamma}^2 + (1-\tilde{\gamma}^2)\gamma \leq \tilde{\gamma}\,. \end{gathered}$$

Hence

$$|||S^{m/2}\tilde{C}S^{m/2}v||| \leq \tilde{\gamma}|||v|||\,.$$

■

The reader may have noticed that it is not essential for $\tilde{C}$ to come from a W-cycle and that could as well come from a V-cycle. There is however a reason to consider the above as a convergence proof for the W-cycle. In a W-cycle the approximation $\tilde{C}$ of C consists of a double application of an approximation $\hat{C}$ of C; $\hat{C}$ coming from a V-cycle. If we now assume that the quality of $\tilde{C}$ as an approximation of C can be expressed as:

$$|||\hat{C}v - Cv||| \leq \tilde{\gamma}|||Cv - v|||,$$

then the quality of $\tilde{C}$ as an approximation of C is given by

$$\begin{aligned} |||\tilde{C}v - Cv||| &= |||\hat{C}\hat{C}v - Cv||| \\ &\leq \tilde{\gamma}|||\hat{C}v - Cv||| \\ &\leq \tilde{\gamma}^2|||v - Cv|||, \end{aligned}$$

which is exactly the main-hypothesis of the previous theorem.

In the following theorem we will make use of the function ϕ which is defined as follows:

Definition 6.25 Let $v = \sum_i c_i\psi_i \in M_j$ and let $S : M_j \to M_j$ be some smoothing operator. Then we define the function

$$\phi : \mathbb{R} \to \mathbb{R}$$

by

$$S^{m/2}v = \sum_i c_i\, \phi(\lambda_i)\, \psi_i\,.$$

Example 6.26 *Jacobi relaxation*
By (6.25) we see that for Jacobi-relaxation we have

$$\phi(\lambda) = (1 - \lambda)^{m/2}\,.$$

Theorem 6.27 Suppose that the hypothesis (6.29) holds. Let γ be as in (6.35). Let $\tilde{C}$ be the approximation of the coarse grid correction operator C satisfying

$$|||\tilde{C}v - Cv||| \leq \hat{\gamma}|||Cv - v||| \text{ for all } v\in M_j\,,$$

where

$$\hat{\gamma} = \frac{1}{1+\beta}, \qquad \beta = \inf_{\lambda\in(0,1]} \frac{\phi^{-2}(\lambda) - 1}{\kappa^\alpha \lambda^\alpha}\,.$$

Then $|||S^{m/2}\tilde{C}S^{m/2}v||| \leq \hat{\gamma}|||v|||$ for all $v\in M_j$.

Proof: Let first $w\in M_j$ be arbitrary. Write $\bar{v} = S^{m/2}v$, $\bar{w} = S^{m/2}w$.

For $T : M_j \to M_j$ a symmetric operator, we have ($a(\cdot,\cdot)$ is symmetric)

$$a(Tu, v) = a(u, Tv).$$

Using this with $T = S^{m/2}$ we find

$$\begin{aligned} a(S^{m/2}\tilde{C}S^{m/2}v, w) &= a(\tilde{C}S^{m/2}v, S^{m/2}w) \\ &= a(\tilde{C}\bar{v}, \bar{w}) \\ &= a(C\bar{v} - \hat{\gamma}C\bar{v} + \tilde{C}\bar{v} - C\bar{v} + \hat{\gamma}C\bar{v}, \bar{w}) \\ &= (1-\hat{\gamma})a(C\bar{v}, \bar{w}) + a((\tilde{C} + (\hat{\gamma} - 1))\bar{v}, \bar{w}) \\ &= (1-\hat{\gamma})a(C\bar{v}, C\bar{w}) + a((\tilde{C} + (\hat{\gamma} - 1))\bar{v}, \bar{w})\,. \end{aligned}$$

The last equality because $C\bar{w}-\bar{w}\in M_{j-1}$ and hence

$$a(C\bar{v},C\bar{w}-\bar{w})=0.$$

Now using (6.23) with $\alpha=0$ gives

$$|a(S^{m/2}\tilde{C}S^{m/2}v,w)| \;\leq\; (1-\hat{\gamma})\ |||C\bar{v}|||\cdot|||C\bar{w}|||+|||\tilde{C}\bar{v}+(\hat{\gamma}-1)C\bar{v}|||\ |||\bar{w}|||\,,$$

or using lemma 6.17

$$|a(S^{m/2}\tilde{C}S^{m/2}v,w)|\leq$$

$$(1-\hat{\gamma})|||C\bar{v}|||\cdot|||C\bar{w}|||+\hat{\gamma}|||\bar{v}|||\cdot|||\bar{w}|||.$$

Now, with

$$\hat{v}=\begin{pmatrix}\sqrt{1-\hat{\gamma}}|||C\bar{v}|||\\ \sqrt{\hat{\gamma}}|||\bar{v}|||\end{pmatrix},\quad \hat{w}=\begin{pmatrix}\sqrt{1-\hat{\gamma}}|||C\bar{w}|||\\ \sqrt{\hat{\gamma}}|||\bar{w}|||\end{pmatrix}$$

Cauchy-Schwartz

$$|<\hat{v},\hat{w}>|\leq\|\hat{v}\|_{\rm eucl}\|\hat{w}\|_{\rm eucl}$$

yields

$$|a(S^{m/2}\tilde{C}S^{m/2}v,w)|\leq$$

$$\sqrt{(1-\hat{\gamma})|||C\bar{v}|||^2+\hat{\gamma}|||\bar{v}|||^2}\cdot\sqrt{(1-\hat{\gamma})|||C\bar{w}|||^2+\hat{\gamma}|||\bar{w}|||^2}.$$

Now we will investigate

$$(1-\hat{\gamma})\ |||C\bar{v}|||^2+\hat{\gamma}\ |||\bar{v}|||^2.$$

Using (6.31) we find

$$(1-\hat{\gamma})\ |||C\bar{v}|||^2+\hat{\gamma}\ |||\bar{v}|||^2\leq(1-\hat{\gamma})\kappa^\alpha\|\bar{v}\|^2_{1+\alpha}+\hat{\gamma}\ |||\bar{v}|||^2\,.$$

Hence $S^{m/2}v=\sum_i\ c_i^v\ \phi(\lambda_i)\psi_i$, and

$$(1-\hat{\gamma})\ |||C\bar{v}|||^2+\hat{\gamma}\ |||\bar{v}|||^2\leq$$

$$\leq\ (1-\hat{\gamma})\kappa^\alpha\sum_i c_i^2\phi^2(\lambda_i)\lambda_i^{1+\alpha}+\hat{\gamma}\sum_i c_i^2\phi^2(\lambda_i)\lambda_i\leq$$

$$\leq\ \max_{\lambda\in[0,1]}[(1-\hat{\gamma})\kappa^\alpha\lambda^\alpha\phi^2(\lambda)+\hat{\gamma}\phi^2(\lambda)]\sum_i c_i^2\lambda_i=$$

$$=\ \max_{\lambda\in[0,1]}[\phi^2(\lambda)((1-\hat{\gamma})\kappa^\alpha\lambda^\alpha+\hat{\gamma})]\cdot|||v|||^2.$$

Returning to the mainline of the proof, substituting the above result, we find

$$|a(S^{m/2}\tilde{C}S^{m/2}v,w|\leq|||v|||\cdot|||w|||\cdot\max_{\lambda\in[0,1]}[\phi^2(\lambda)((1-\hat{\gamma})\kappa^\alpha\lambda^\alpha+\hat{\gamma}].$$

Taking $w=S^{m/2}\tilde{C}S^{m/2}v$ and knowing that

$$a(S^{m/2}\tilde{C}S^{m/2}v,S^{m/2}\tilde{C}S^{m/2}v)=|||S^{m/2}\tilde{C}S^{m/2}v|||^2,$$

division by $|||S^{m/2}\tilde{C}S^{m/2}v|||$ yields

$$|||S^{m/2}\tilde{C}S^{m/2}v|||\leq\max_{\lambda\in[0,1]}[\phi^2(\lambda)((1-\hat{\gamma})\kappa^\alpha\lambda^\alpha+\hat{\gamma})]\ |||v|||\,. \tag{6.37}$$

Let $\beta = \inf_{\lambda\in[0,1]} \frac{\phi^{-2}(\lambda)-1}{\kappa^\alpha\lambda^\alpha} \implies \beta\kappa^\alpha\lambda^\alpha+1 \leq \phi^{-2}(\lambda)$ for all $\lambda\in[0,1] \implies \phi^2(\lambda)(\beta\kappa^\alpha\lambda^\alpha+1)\leq 1$ for all $\lambda\in[0,1]$. Now $\hat\gamma = \frac{1}{1+\beta} \implies \beta = \frac{1-\hat\gamma}{\hat\gamma}$. Substitution of $\frac{1-\hat\gamma}{\hat\gamma}$ for β gives

$$\phi^2(\lambda)((\frac{1-\hat\gamma}{\hat\gamma})\kappa^\alpha\lambda^\alpha+1)\leq 1 \text{ for all } \lambda\in[0,1] \implies$$

$$\phi^2(\lambda)((1-\hat\gamma)\kappa^\alpha\lambda^\alpha+\hat\gamma)\leq \hat\gamma \text{ for all } \lambda\in[0,1] \implies$$
$$\max_{\lambda\in[0,1]}[\phi^2(\lambda)((1-\hat\gamma)\kappa^\alpha\lambda^\alpha+\hat\gamma)] \leq \hat\gamma\,.$$

Now the proof follows by substitution of this inequality in (6.37). ∎

Example 6.28 *Jacobi relaxation*
In this example we want to compute $\hat\gamma$ when Jacobi-relaxation is used. From example 6.26 we know $\phi(\lambda) = (1-\lambda)^{m/2}$.

Hence
$$\beta = \inf_{\lambda\in[0,1]} \frac{(1-\lambda)^{-m}-1}{\kappa^\alpha\lambda^\alpha} = \frac{1}{\kappa^\alpha}\inf_{\lambda\in[0,1]}\frac{(1-\lambda)^{-m}-1}{\lambda^\alpha}\,.$$

For simplicity we assume $\alpha\in(0,1]$.

Define
$$f(\lambda) = \frac{(1-\lambda)^{-m}-1}{\lambda^\alpha} \quad \text{on} \quad (0,1)\,,$$

then
$$f'(\lambda) = \frac{(1-\lambda)^{m-1}}{\lambda^{\alpha+1}}\cdot(m\lambda-\alpha(1-\lambda)+\alpha(1-\lambda)^{m+1}).$$

We see that $f'(\lambda)\geq 0$ on (0,1) if

$$g(\lambda) = m\lambda-\alpha(1-\lambda)+\alpha(1-\lambda)^{m+1}\geq 0 \quad \text{on} \quad (0,1)\,,$$

and
$$\begin{aligned} g'(\lambda) &= m+\alpha-(m+1)\alpha(1-\lambda)^m \\ &\geq m+\alpha-(m+1)\alpha \geq 0 \;\text{ since }\; \alpha\in(0,1]. \end{aligned}$$

Further $g(0)=0$ and hence $g(\lambda)\geq 0$ on $[0,1]$. We now conclude that $f'(\lambda)\geq 0$ on [0,1]. Using this we find

$$\beta = \frac{1}{\kappa^\alpha}\lim_{\lambda\to 0}\frac{(1-\lambda)^{-m}-1}{\lambda^\alpha} = \begin{cases} \frac{m}{\kappa}, & \alpha = 1 \qquad \text{(Taylor)} \\ 0, & 0<\alpha<1 \end{cases}$$

Hence
$$\hat\gamma = \frac{1}{1+\beta} = \begin{cases} \frac{\kappa}{\kappa+m}, & \text{if } \alpha = 1, \\ 1, & \text{if } 0<\alpha<1. \end{cases}$$

Chapter 7

Local mode analysis

7.1 Fourier transforms of continuous functions

In this section we collect well-known results with respect to Fourier transforms of functions that are defined (almost everywhere) on domains in the real n-dimensional space. All results mentioned in this section can be found in general texts as e.g. [14], [17], [19], [20].

Let $u \in L^2(\mathbb{R}^n)$, then its Fourier transform $\widehat{u}$ is defined by

$$\widehat{u}(y) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ixy}\, u(x)\, dx\,. \tag{7.1}$$

Furthermore, a back-transformation formula is available:

$$\tilde{u}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{+ixy}\, \widehat{u}(y)\, dy\,, \tag{7.2}$$

such that $\tilde{u}(x) = u(x)$ almost everywhere on $\mathbb{R}^n$. Moreover

$$\widehat{u} \in L^2(\mathbb{R}^n) \quad \text{and} \quad \|u\|_{L^2(\mathbb{R}^n)} = \|\widehat{u}\|_{L^2(\mathbb{R}^n)}\,.$$

We say that the Fourier transformation is a norm-invariant bijection $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

The above definition of a Fourier transformation can be generalised to more general functions than $L^2(\mathbb{R}^n)$-functions. The same definition applies to the set of "tempered distributions" (see e.g. [20]); in this case -again- the back transformation is available.

We see that the Fourier transform (FT) of a function defined on $\mathbb{R}^n$ is a function defined on $\mathbb{R}^n$ itself. A Fourier transformation can also be defined for a finite set of equally spaced data. In this case the FT of a set of N data (the Finite Fourier transform) is again a set of N coefficients (see e.g. [11]).

The FT of a periodic function (or, what is the same, the FT of a function defined on a torus) is a countable infinite set of coefficients. Analogously, in the following sections we shall introduce the Fourier transformation on an infinite set of equally spaced data. In this case the FT of such a "gridfunction" will be periodic function (a function defined on a torus).

Definition 7.1 Let $h = (h_1, ..., h_n) \in \mathbb{R}^n$ be given, then the *h-periodisation* of a function $u : \mathbb{R}^n \to \mathbb{C}$ is defined by

$$\tilde{u}(x) = \sum_{k \in \mathbb{Z}^n} u(x - kh),$$

where $kh = (k_1 h_1, ..., k_n h_n)$.

We notice that $\tilde{u}(x)$ is a periodic function on $\mathbb{R}^n$ with period h; it is completely defined by a mapping $[0,h) \to \mathbb{C}$, where $[0,h)$ is defined by

$$[0,h) = [0,h_1) \times [0,h_2) \times ... \times [0,h_n)\,.$$

The FT of a function $\tilde{u}(x)$ defined on the torus $[0,h)$ is (cf. [14]) a sequence $\{c_k\}_{k\in\mathbb{Z}^n}$ defined by

$$c_k = \frac{1}{h^n(2\pi)^n}\int_0^h e^{-2\pi ikx/h}\tilde{u}(x)dx, \tag{7.3}$$

from which it is clear that $c_k = \widehat{u}(\frac{2\pi k}{h})/h^n$. Also the Fourier transform on $[0,h)$ has its back-transformation. From this we see that the knowledge of $\widehat{u}(y)$ at only certain equally spaced points is enough to restore a periodisation of the original function u, whereas the complete definition of $\widehat{u}(y)$ (almost everywhere on $\mathbb{R}^n$) is needed to find the function $u(x)$ itself.

7.2 Gridfunctions

For a fixed "mesh" $h = (h_1,...,h_n)$ with $h_i > 0$, $i = 1,2,...,n$, the regular infinite n-dimensional grid $\mathbb{Z}_h^n$ is defined by

$$\mathbb{Z}_h^n = \{jh | j\in\mathbb{Z}^n\}\,.$$

Further we introduce the notation

$$j/h = (j_1/h_1,...,j_n/h_n),$$

$$h^n = h_1 \cdot h_2 \cdot ... \cdot h_n,$$

$$T_h^n = (2\pi/h]^n = (-\pi/h_1, \pi/h_1] \times ... \times (-\pi/h_n, \pi/h_n].$$

We call T_h^n an n-dimensional torus.

Definition 7.2 A complex or a real *gridfunction* is a mapping $\mathbb{Z}_h^n \to \mathbb{C}^d$, respectively $\mathbb{Z}_h^n \to \mathbb{R}^d$, where d is the dimension of the range of the mapping.

Remark: We will restrict ourselves to the scalar real or complex gridfunction $\mathbb{Z}_h^n \to \mathbb{R}$ or $\mathbb{Z}_h^n \to \mathbb{C}$ and, unless stated otherwise, we shall use the word gridfunction for this kind of gridfunction exclusively. It is immediate that, with the usual addition and scalar multiplication, the set of all gridfunctions is a vector space. This vector space we denote by

$$l_h(\mathbb{Z}_h^n)$$

or, shortly, by l_h. For any $p \geq 1$ or $p = \infty$ the space l_h can be provided with a norm

$$\|u_h\|_p = (h^n \sum_{j\in\mathbb{Z}^n} |u_h(jh)|^p)^{1/p}, \quad 1 \leq p < \infty, \tag{7.4}$$

or

$$\|u_h\|_\infty = \ \sup_{j\in\mathbb{Z}^n}|u_h(jh)|\,. \tag{7.5}$$

For a fixed p, $1 \leq p \leq \infty$, all gridfunctions with a finite norm $\|\cdot\|_p$ form a subspace of $l_h(\mathbb{Z}_h^n)$, which is denoted by

$$l_h^p(\mathbb{Z}_h^n)\,.$$

It is obvious that for any p, $1 \leq p \leq \infty$, $l_h^p(\mathbb{Z}_h^n)$ is a Banach space (cf. [25] p.35). Moreover, for $p = 2$ we know that $l_h^2(\mathbb{Z}_h^n)$ is a Hilbert space with the inner product

$$< u_h, v_h >_h = h^n \sum_{j\in\mathbb{Z}^n} u_h(jh)\overline{v_h(jh)}\,. \tag{7.6}$$

7.3 The Fourier transform of a gridfunction

Let $u_h : \mathbb{Z}_h^n \to \mathbb{C}$ be a gridfunction. We give the following

Definition 7.3 The Fourier transform $\widehat{u_h} \in L^2(T_h^n)$ of $u_h \in l_h^2(\mathbb{Z}_h^n)$ is a function $T_h^n \to \mathbb{C}$, defined by

$$\widehat{u_h}(\omega) = \left(\frac{h}{\sqrt{2\pi}}\right)^n \sum_{j\in\mathbb{Z}^n} e^{-ijh\omega}\, u_h(jh)\,. \tag{7.7}$$

The inverse transformation is given by

$$u_h(jh) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int_{\omega\in T_h^n} e^{+ijh\omega}\, \widehat{u_h}(\omega)\, d\omega\,. \tag{7.8}$$

Remarks:

- We denote the Fourier transform also by

$$\widehat{u_h} \quad = \quad \mathbf{FT}(u_h).$$

- $\widehat{u_h}$ can also be considered as a $[2\pi/h]^n$-periodic function $\widehat{u_h} : \mathbb{R}^n \to \mathbb{C}$.
- By the *Parseval identity* we have

$$\|u_h\|_2 = \|\widehat{u_h}\|_{L^2(T_h^n)}\,.$$

  Hence the Fourier transformation operator $\mathbf{FT}$: $l_h^2 \to L^2(T_h^n)$ is a unitary operator.
- With the identity $\int_{T_h^n} e^{ih\omega(j-k)} d\omega = (\frac{2\pi}{h})^n \delta_{jk}$ we easily verify the back-transformation formula.

In the back-transformation formula (7.8) we see that any gridfunction $u_h \in l_h^2$ can be considered as a linear combination of gridfunctions $e_{h,\omega}$ of the form

$$e_{h,\omega}(jh) = e^{ijh\omega}, \quad \omega\in T_h^n,\, j\in\mathbb{Z}^n, \tag{7.9}$$

i.e. a periodic gridfunction with period $[\frac{2\pi}{h\omega}]^n$; such $e_{h,\omega}$ is called a *simple mode* . The parameter ω is called the *frequency* of this mode and $\widehat{u_h}(\omega)$ is the *amplitude* of the component (mode) with frequency ω in u_h.

Because $e_{h,\omega} \equiv e_{h,\omega+2\pi k/h}$, for all $k\in\mathbb{Z}^n$, a frequency ω cannot be distinguished from a frequency $\omega + 2\pi k/h$. This phenomenon is called *aliasing* .

It is clear from definition 7.3 that the range of frequencies that can be represented on a fine grid (small h) is larger than the range of those which can be represented on a coarser grid (large h). This explains why smooth gridfunctions (i.e. gridfunctions with relatively small amplitudes for the high frequencies) can be much better represented on a coarser grid than less smooth gridfunctions.

Remarks:

- For functions u_h defined on a bounded domain Ω_h, $\Omega_h = [-Nh, Nh]^n \bigcap \mathbb{Z}_h^n$, with periodic boundary conditions, a similar definition can be given for the Fourier transform $\widehat{u_h}$. The only difference is that the set of frequencies, T_h^n, available for such gridfunction u_h is discrete:

$$T_h^n = \{(\pi k_1/Nh, ..., \pi k_n/Nh) \mid -N < k_j < N,\ j = 1, ..., n\} \tag{7.10}$$

- Functions on a bounded domain with homogeneous Dirichlet or Neumann boundary conditions can be considered as a special case of problems with periodic boundary conditions. In this case the elementary components in these functions are all of the form
$$e_{h,\omega} \pm e_{h,\omega} = e^{ijh\omega} \pm e^{-ijh\omega}\,.$$
and the Fourier analysis method gives an exact description of symmetric operators B_h : $l_h(\Omega_h) \to l_h(\Omega_h)$ because for symmetric operators we have $(B_h)_{i,k} = (B_h)_{k,i}$ and hence $\widehat{B_h}(\omega) = \widehat{B_h}(-\omega)$. (See Section 7.5 for a definition of $\widehat{B_h}$).

7.4 The relation between FTs of a function restricted to different grids

In this section we will present two theorems. One gives the relation between the FT of a continuous function defined on $\mathbb{R}^n$ and the FT of its restriction to the grid $\mathbb{Z}^n_h$. The other gives the relation between the FT of a gridfunction and the FT of its canonical restriction to a coarser grid.

Theorem 7.4 Let $u \in L^2(\mathbb{R}^n)$ be a continuous function with FT $\widehat{u}$. Its restriction to the grid $\mathbb{Z}^n_h$ is defined by
$$u_h(jh) = (R^0_h u)(jh) = u(jh),\; j \in \mathbb{Z}^n\,. \tag{7.11}$$
Then we have the following relation between $\widehat{u}$ and $\widehat{u_h}$:
$$\widehat{u_h}(\omega) = \sum_{k \in \mathbb{Z}^n} \widehat{u}(\omega + 2\pi k/h), \tag{7.12}$$
i.e. $\widehat{u_h}$ is the $[2\pi/h]^n$-periodisation of $\widehat{u}$ (def. 7.1).
Proof:
$$\begin{aligned}
\widehat{u_h}(\omega) &= (\tfrac{h}{\sqrt{2\pi}})^n \textstyle\sum_j e^{-ijh\omega} u(jh)\\
&= (\tfrac{h}{2\pi})^n \textstyle\sum_j e^{-ijh\omega} \sum_k \int_{-\pi/h}^{\pi/h} e^{ijh(y+2\pi k/h)}\, \widehat{u}(y + 2\pi k/h) dy\\
&= (\tfrac{h}{2\pi})^n \textstyle\sum_j e^{-ijh\omega} \int_{-\pi/h}^{\pi/h} e^{ijhy} \sum_k \widehat{u}(y + 2\pi k/h) dy
\end{aligned}$$
Using (7.7) and (7.8) we see that this equals
$$\sum_{k \in \mathbb{Z}^n} \widehat{u}(\omega + 2\pi k/h)\,.$$
■

Definition 7.5 Let $u_h \in l_h(\mathbb{Z}^n_h)$ then its *canonical q-restriction* $R^0_q u_h, (q \in \mathbb{Z}^n)$, is the gridfunction u_H defined on $\mathbb{Z}^n_H = \mathbb{Z}^n_{qh}$, defined by
$$(R^0_q u_h)(jH) = u_H(jH) = u_h(jqh)\,. \tag{7.13}$$
Sometimes we write R^0_{Hh}, with $H = qh$, instead of R^0_q.

Lemma 7.6 If $u_h \in l^p_h(\mathbb{Z}^n_h)$ then $R^0_q u_h \in l^p_H(\mathbb{Z}^n_H)$ with $H = qh$.
Proof:
$$\begin{aligned}
\|u_H\|^p_{l^p_H(\mathbb{Z}^n_H)} &= q^n h^n \textstyle\sum_{j\in\mathbb{Z}^n} |u_h(jqh)|^p\\
&\le q^n h^n \textstyle\sum_{j\in\mathbb{Z}^n} |u_h(jh)|^p\\
&\le q^n \|u_h\|^p_{l^p_h(\mathbb{Z}^n_h)} < \infty\,.
\end{aligned}$$
■

Theorem 7.7

$$(\widehat{R_q u_h})(\omega) = \sum_{p\in[0,q)} \widehat{u_h}(\omega + 2\pi p/H), \text{ for all } \omega\in T_H^n, H = qh, q\in\mathbb{Z}^n, q > 0\,. \tag{7.14}$$

Proof: The proof is left as an exercise. ■

Theorem 7.7 shows us that, using the restriction R_q^0 with $q\in\mathbb{Z}^n, q = (q_1, ..., q_n)$, we get aliasing of $Q = q_1 ....q_n$ frequencies onto one.

7.5 Toeplitz operators and their FTs

Let $A : l_h(\mathbb{Z}_h^n) \to l_h(\mathbb{Z}_h^n)$ be a linear operator. We are interested to know if for any given $\omega\in T_h^n$ the operator A has an eigenvalue λ_ω corresponding with an eigenfunction $e_{h,\omega}$ as defined by (7.9). Suppose it to be true. Then, denoting A by its matrix representation (a_{mj}), $m, j\in\mathbb{Z}^n$, the following holds

$$\sum_{j\in\mathbb{Z}^n} a_{mj} e_{h,\omega}(jh) = \lambda_\omega e_{h,\omega}(mh)$$

$$\sum_{j\in\mathbb{Z}^n} a_{mj} e^{ijh\omega} = \lambda_\omega e^{imh\omega},$$

hence

$$\lambda_\omega = \sum_{j\in\mathbb{Z}^n} a_{mj} e^{i(j-m)h\omega} = \sum_{k\in\mathbb{Z}^n} a_{m,m+k} e^{ikh\omega}\,.$$

Because λ_ω is independent of m, an a_{-k} should exist such that $a_{m,m+k} = a_{-k}$ for all $m\in\mathbb{Z}^n$. These considerations give rise to the following

Definition 7.8 A linear operator $A : l_h(\mathbb{Z}_h^n) \to l_h(\mathbb{Z}_h^n)$ whose matrix elements a_{mj} satisfy the relation $a_{m,m+k} = a_{-k}$ for all $m\in\mathbb{Z}^n$ and certain a_{-k}, is called a *Toeplitz operator*.

Lemma 7.9 Let $A : l_h(\mathbb{Z}_h^n) \to l_h(\mathbb{Z}_h^n)$ be a Toeplitz operator with matrix representation (a_{mj}). Then for any $\omega\in T_h^n$

$$\lambda_\omega = \sum_{k\in\mathbb{Z}^n} a_{-k} e^{ikh\omega} \tag{7.15}$$

is an eigenvalue of A, corresponding with the eigenfunction $e_{h,\omega}$.

Definition 7.10 Let $a_h, u_h\in l_h(\mathbb{Z}_h^n)$ be two gridfunctions. The a_h*-convolution* of u_h, denoted by $a_h \star u_h\in l_h(\mathbb{Z}_h^n)$, is defined by

$$(a_h \star u_h)(mh) = \sum_{j\in\mathbb{Z}^n} a_h((m-j)h)\ u_h(jh) \text{ for all } m\in\mathbb{Z}^n. \tag{7.16}$$

We can identify the a_h-convolution with a Toeplitz operator A_h on $l_h(\mathbb{Z}_h^n)$ in the following sense:

$$A_h : l_h(\mathbb{Z}_h^n) \to l_h(\mathbb{Z}_h^n) \text{ such that } : A_h u_h = a_h \star u_h. \tag{7.17}$$

I.e. the Toeplitz operator A_h is uniquely identified by the gridfunction $a_h\in l_n(\mathbb{Z}_h^n)$ as an a_h-convolution. The gridfunction a_h is related to the matrix elements of A by the relation

$$a_h(-kh) = a_{-k} \text{ for all } k\in\mathbb{Z}^n\,. \tag{7.18}$$

By (7.18) follows

$$\begin{array}{rl} (Au_h)(mh) & = \sum_{j\in \mathbb{Z}^n} a_{mj}\, u_h(jh) \\ & = \sum_{j\in \mathbb{Z}^n} a_{-(j-m)}\, u_h(jh) \\ & = \sum_{j\in \mathbb{Z}^n} a_h((m-j)h)\, u_h(jh) \\ & = (a_h * u_h)(mh), \quad \text{for all} \quad m\in\mathbb{Z}^n,\ u_h\in l_h. \end{array} \tag{7.19}$$

We say that the gridfunction a_h *generates* the Toeplitz operator $A = a_h \star$.

Definition 7.11 The *Fourier transform* $\widehat{A_h} : T_h^n \to \mathbb{C}$ of a *Toeplitz operator* A_h is defined by

$$A_h e_{h,\omega} = \widehat{A_h}(\omega) e_{h,\omega}\,. \tag{7.20}$$

Using (7.15) we see that this means that

$$\widehat{A_h}(\omega) = \lambda_\omega = \sum_{k\in\mathbb{Z}^n} a_h(kh) e^{-ikh\omega}. \tag{7.21}$$

An immediate consequence of this is the following

Lemma 7.12

$$\widehat{A_h u_h}(\omega) = \widehat{A_h}(\omega)\widehat{u_h}(\omega). \tag{7.22}$$

Combining (7.21) and definition 7.3 we obtain the following

Lemma 7.13

$$\widehat{A_h}(\omega) = \left(\frac{\sqrt{2\pi}}{h}\right)^n \widehat{a_h}(\omega)\,.$$

Examples of Toeplitz operators are linear difference operators with constant coefficients, defined on the space of gridfunctions, such that the same difference equation is applied at each gridpoint (i.e. the matrix elements of this operator satisfy $a_{m,m+k} = a_{-k}$ for all $m\in\mathbb{Z}^n$). Another example is the translation operator T_{kh}

$$(T_{kh}u_h)(jh) = u_h((j-k)h)\,. \tag{7.23}$$

A typical property of, for example, the above mentioned linear difference operators, is that the difference equation applied to some gridpoint only relates a few neighbouring gridpoints. Examining the definition of a convolution product

$$(a_h \star u_h)(mh) = \sum_{k\in\mathbb{Z}^n} a_h(kh)\ u_h((m-k)h)$$

we see that this means that the generating gridfunction a_h has *finite support:* $a_h(kh) \neq 0$ only for some $k\in\mathbb{Z}^n$ in a neighbourhood of $0\in\mathbb{Z}^n$.

Suppose now that the Toeplitz operator A_h is generated by a gridfunction a_h with finite support, i.e.

$$\left\{ \begin{array}{ll} a_h(-kh) = s_k & k\in V, \\ a_h(-kh) = 0 & k\notin V, \end{array} \right.$$

for some *finite* set $V \subset \mathbb{Z}^n$, containing $0\in\mathbb{Z}^n$. Then we often associate A_h with a *stencil.* For $n = 2$ such a stencil is given by

$$A_h = \begin{bmatrix} & \vdots & \vdots & \vdots & \\ \cdots & s_{-1,1} & s_{0,1} & s_{1,1} & \cdots \\ \cdots & s_{-1,0} & s_{0,0} & s_{1,0} & \cdots \\ \cdots & s_{-1,-1} & s_{0,-1} & s_{1,-1} & \cdots \\ & \vdots & \vdots & \vdots & \end{bmatrix}_h ,$$

where only non-zero s_k are given. This notation (the *stencil notation*) is very convenient if V is small.

We will now give a few detailed examples of Toeplitz operators and their FT.

Example 7.14
Consider a central difference discretisation of the Laplace operator Δ in one dimension on a regular infinite grid $\mathbb{Z}_h^1$. The matrix representation of the resulting discrete operator L_h is given in Chapter 4. In many cases the stencil representation is more convenient:

$$L_h \;=\; \frac{1}{h^2}\,[1,-2,1]_h\,.$$

We can now, very easily, determine $\widehat{L_h}$, using (7.21):

$$\widehat{L_h}(\omega) = \frac{1}{h^2}(e^{ih\omega} \;+\; e^{-ih\omega} \;-\; 2) = \frac{1}{h^2}(2\cos(h\omega) \;-\; 2) = \;-\frac{4}{h^2}\sin^2(h\omega/2)\,. \qquad (7.24)$$

Example 7.15
The translation operator T_{kh} (see (7.23)) can be written as a convolution:

$$(T_{kh}u_h)(mh) = u_h((m-k)h) = \sum_{j\in\mathbb{Z}^n} a_h(jh)\,u_h((m-j)h),$$

where $a_h(kh) = 1$, and $a_h(jh) = 0$ if $k \neq j$. Hence we can identify T_{kh} with the following stencil ($n = 1$):

$$T_{kh} = [1,0,...,0,\underline{0},0,...,0].$$

with k terms on either side of $\underline{0}$.
Its FT is given by

$$\widehat{T_{kh}}(\omega) = e^{-ikh\omega} = \cos(kh\omega) - i\,\sin(kh\omega)\,.$$

Because $|\widehat{T_{kh}}(\omega)| = 1$ for all $\omega\in T_h^1$ we see that application of a translation operators to u_h does not affect the absolute value of $\widehat{u_h}(\omega)$.

7.6 Consistency of a discrete operator

For linear differential operators with constant coefficients we can use Fourier analysis to determine the order of consistency of a discretisation of this operator. First we recall the following, well known property . A linear partial differential operator L of order N with constant coefficients, given by

$$L = \sum_{|\alpha|\le N} c_\alpha \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}...\partial x_n^{\alpha_n}}\,, \quad (|\alpha| = \alpha_1 \;+\; ... \;+\; \alpha_n)\,,$$

has the property that for any $\xi\in\mathbb{C}^n$ the function $\mathbb{R}^n \to \mathbb{C}:\; x \to e^{ix\cdot\xi}$ is mapped by L onto a multiple of itself:

$$Le^{ix\cdot\xi} = p(\xi)e^{ix\cdot\xi},$$

with $p(\xi) = \sum_{|\alpha|\le N} c_\alpha(i\xi)^\alpha$, the characteristic polynomial. Here $x\cdot\xi = \sum_{k=1}^n x_k\xi_k$ and $\eta^\alpha = \eta_1^{\alpha_1}...\eta_n^{\alpha_n}$. The characteristic polynomial p is called the *symbol* of L. In analogy to definition 7.11 we can write

$$\widehat{L}(\omega) = p(\omega)\,.$$

Consider a linear partial differential operator L with constant coefficients and its discretisation L_h on some grid $\mathbb{Z}_h^n$. We are interested in the truncation error

$$\tau_h = L_hR_h \;-\; \bar{R}_hL, \quad \tau_h : E \to \hat{E}_h$$

(see Figure 2.1 for a summary of the operators and notation used). R_h and $\bar{R}_h$ are the canonical restriction. For this purpose we consider some arbitrary $\omega \in T_h^n$ and the truncation error of L_h for the function

$$e_\omega(x) = e^{i\omega\cdot x}, \quad e_\omega \in E\,.$$

We find

$$\begin{aligned}\tau_h(e_\omega)(x) &= (L_h R_h e_\omega - \bar{R}_h L e_\omega)(x)\\ &= L_h e^{i\omega jh} - R_h \widehat{L}(\omega) e^{i\omega\cdot x}\\ &= (\widehat{L_h}(\omega) - \widehat{L}(\omega))e^{ijh\omega},\end{aligned}$$

hence

$$\|\tau_h(e_\omega)\| = |\widehat{L_h}(\omega) - \widehat{L}(\omega)|. \tag{7.25}$$

For given L and L_h this expression can be developed in powers of h and the order of consistency of L_h can be determined.

Example 7.16 *Consistency order*
We take $L = \Delta$ (in two dimensions),

$$L_h = \frac{1}{h^2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}_h .$$

We find

$$\begin{aligned}\widehat{L}(\omega) &= -\omega_1^2 - \omega_2^2,\\ \widehat{L_h}(\omega) &= \tfrac{1}{h^2}[2\cos(h\omega_1) + 2\cos(h\omega_2) - 4]\\ &= -\tfrac{1}{h^2}[4\sin^2(\omega_1 h/2) + 4\sin^2(\omega_2 h/2)].\end{aligned}$$

Hence

$$\begin{aligned}|\widehat{L_h}(\omega) - \widehat{L}(\omega)| &= |\frac{4}{h^2}\sin^2(\omega_1 h/2) + \frac{4}{h^2}\sin^2(\omega_2 h/2) - \omega_1^2 - \omega_2^2|\\ &= \frac{1}{12}h^2\,|\omega_1^4 + \omega_2^4| + \mathcal{O}(h^4) \quad \text{for} \quad h\to 0\,.\end{aligned}$$

We conclude that the order of consistency of L_h is 2.

7.7 The smoothing factor for relaxation methods

Consider a multigrid algorithm where for each coarser mesh the meshwidth is related to finer mesh by $H = 2h$. In Section 7.3 we saw that the range of frequencies that can be represented on the coarser grid is smaller than the range available on the finer grid. For $H = 2h$ the frequencies that cannot be correctly transferred (i.e. without aliasing) to the next coarser grid are all $\omega \in T_h^n \setminus T_{2h}^n$.

If we want the coarse-grid-correction in the multigrid algorithm to be successful, the relaxation steps should damp the amplitudes of the error modes with frequencies $\omega \in T_h^n \setminus T_{2h}^n$. This smoothing property of a given relaxation method with amplification operator M_h^{REL} for the error, is expressed in terms of the so called *smoothing factor* μ:

$$\mu = \sup_{\omega\in T_h^n \setminus T_{2h}^n} |\widehat{M_h^{REL}}(\omega)|\,. \tag{7.26}$$

This smoothing factor μ describes how well the high-frequency components in the error are reduced by the relaxation method.

If L_h is a Toeplitz operator and B_h, an approximate inverse of L_h, is a Toeplitz operator as well, then

$$M_h^{REL} = I_h - B_h L_h$$

is a Toeplitz operator, and

$$\widehat{M_h^{REL}}(\omega) = I - \widehat{B_h}(\omega)\widehat{L_h}(\omega)\,. \tag{7.27}$$

Example 7.17 *computation smoothing factor*
We return to Example 7.14 where we considered

$$L_h = \frac{1}{h^2}[1,-2,1]_h, \qquad \widehat{L_h}(\omega) = \frac{1}{h^2}(2\cos(h\omega)-2).$$

As a relaxation method we choose damped Jacobi relaxation with parameter α, i.e. we use

$$B_h = \alpha(\mathrm{diag}(L_h))^{-1} = h^2\,[0, \frac{-\alpha}{2}, 0]\,.$$

We find $\widehat{B_h}(\omega) = \frac{-h^2\alpha}{2}$, hence

$$\widehat{M_h^{REL}}(\omega) = 1 - \widehat{B_h}(\omega)\widehat{L_h}(\omega) = 1 - \alpha + \alpha\cos(h\omega)\,.$$

One can show that $\alpha = 2/3$ yields optimal smoothing of this relaxation method for the given problem. For $\alpha = 2/3$ the smoothing factor is $\mu = 1/3$.

7.8 Restrictions and prolongations

In section 7.4 we introduced the canonical restriction R_h^0 and the canonical q-restrictions $R_q^0 : l_h(\mathbb{Z}_h^n) \to l_{qh}(\mathbb{Z}_{qh}^n)$. Now we show that an arbitrary restriction, where for each gridpoint the same stencil is used) can be described in terms of a Toeplitz operator.

Definition 7.18 A homogeneous (*weighted*) *q-restriction* of a gridfunction u_h defined on $\mathbb{Z}_h^n$ to a gridfunction $u_H, H = qh$, defined on $\mathbb{Z}_H^n$, denoted by $R_q(a_h)u_h$, is defined by

$$(R_q(a_h)u_h)(jH) = u_H(jH) = \sum_{k\in\mathbb{Z}^n} a_h(kh)\; u_h((qj-k)h)\,. \tag{7.28}$$

The gridfunction a_h generates a Toeplitz operator. If we denote this operator by A_h (i.e. $A_h = a_h\star$) the expression (7.28) can be written in the following form:

$$(R_q(a_h)u_h)(jH) = (R_q^0 A_h u_h)(jqh)\,. \tag{7.29}$$

From this we may conclude, using theorem 7.7 and lemma 7.12, that

$$\mathbf{FT}(R_q(a_h)u_h)(\omega) = \sum_{p\in[0,q)} \widehat{A_h}(\omega + 2\pi p/qh)\;\widehat{u_h}(\omega+2\pi p/qh) \tag{7.30}$$

The frequencies $\omega + 2\pi p/qh$, $p \neq 0$, we call the *higher harmonics* of the frequency ω. The total number of harmonics (ω itself included) is

$$Q = q_1 .....q_n.$$

Let us first consider the case $n = 1, q = 2$. Then we have

$$\mathbf{FT}(R_q(a_h)u_h)(\omega) = \sum_{p=0}^{1} \widehat{A_h}(\omega + \pi p/h)\,\widehat{u_h}(\omega + \pi p/h).$$

Another notation of this is

$$\widehat{u_H}(\omega) = \mathbf{FT}(R_q(a_h)u_h)(\omega) = (\widehat{A_h}(\omega),\ \widehat{A_h}(\omega+\pi/h)) \begin{pmatrix} \widehat{u_h}(\omega) \\ \widehat{u_h}(\omega+\pi/h) \end{pmatrix}$$

$$=: \vec{\mathbf{FT}}(R_q(a_h))(\omega) \cdot \vec{\mathbf{FT}}(u_h)(\omega), \quad \omega \in T^1_{qh}.$$

For general $n = 1, q > 1, \omega \in T^n_{qh}$, we write also

$$\mathbf{FT}(R_q(a_h)u_h)(\omega) = \vec{\mathbf{FT}}(R_q(a_h))(\omega)\ \vec{\mathbf{FT}}(u_h)(\omega), \tag{7.31}$$

where $\vec{\mathbf{FT}}(R_q(a_h))(\omega)$ is a $1 \times Q$-matrix with components $\widehat{A_h}(\omega + 2\pi p/qh)$, $p \in [0, q)$, and $\vec{\mathbf{FT}}(u_h)(\omega)$ a vector with Q components $\widehat{u_h}(\omega + 2\pi p/qh)$, $p \in [0, q)$.

Example 7.19 $(n = 1, q = 2)$

$$A_h = a_h\star = [1/4, 1/2, 1/4]_h,$$

$$\widehat{A_h}(\omega) = \frac{1}{2} + \frac{1}{2}\cos(\omega h) = \cos^2(\omega h/2),$$

$$\widehat{A_h}(\omega + \pi/h) = \sin^2(\omega h/2),$$

$$\mathbf{FT}(R_q(a_h))(\omega) = (\cos^2(\omega h/2)\,, \sin^2(\omega h/2))\cdot$$

Example 7.20 $(n = 2, q = 2)$

$$A_h = a_h\star = \begin{bmatrix} & 1/8 & \\ 1/8 & 1/2 & 1/8 \\ & 1/8 & \end{bmatrix}_h ,$$

Here $\widehat{A_h}(\omega) = \frac{1}{2}\cos^2(\omega_1 h/2) + \frac{1}{2}\cos^2(\omega_2 h/2)$, and analogously we find $\widehat{A_h}(\omega_1,\ \omega_2 + \pi/h)$, $\widehat{A_h}(\omega_1 + \pi/h,\ \omega_2)$ and $\widehat{A_h}(\omega_1 + \pi/h,\ \omega_2 + \pi/h)$. Thus, we find

$$\mathbf{FT}(R_q(a_h))(\omega) = \frac{1}{2} \begin{pmatrix} \cos^2(\omega_1 h/2) + \cos^2(\omega_2 h/2) \\ \cos^2(\omega_1 h/2) + \sin^2(\omega_2 h/2) \\ \sin^2(\omega_1 h/2) + \cos^2(\omega_2 h/2) \\ \sin^2(\omega_1 h/2) + \sin^2(\omega_2 h/2) \end{pmatrix}.$$

Definition 7.21 Let $u_H \in l_H(\mathbb{Z}^n_H)$ be a gridfunction defined on $\mathbb{Z}^n_H$, then its *flat q-prolongation* $P^0_q u_H$, $q \in \mathbb{Z}^n$, is the gridfunction u_h, $h = H/q$, $u_h \in l_h(\mathbb{Z}^n_h)$, defined by

$$u_h(jh) = (P^0_q u_H)(jh) = \begin{cases} u_H(kqh), & \text{if } j = kq, \\ 0 & \text{otherwise.} \end{cases} \tag{7.32}$$

One can easily verify the following

Theorem 7.22

$$\mathbf{FT}(P^0_q u_H)(\omega) = \widehat{u_h}(\omega) = q^{-n}\,\widehat{u_H}(\omega),\ \omega \in T^n_h\cdot \tag{7.33}$$

Proof: The proof is left as en exercise. ∎

The interpretation of this theorem is that $\widehat{u_h}(\omega)$ is the periodic continuation of $\widehat{u_H}(\omega)$, except for the factor

$$q^{-n} = Q^{-1}.$$

Analogous to definition 7.18 any homogeneous q-prolongation can be described in terms of a Toeplitz operator $A_h = a_h\star$:

$$P_q(a_h) = A_h P_q^0. \tag{7.34}$$

Example 7.23 *Linear interpolation in one dimension ($n = 1, q = 2$)*

$$A_h = a_h\star = [\frac{1}{2}, 1, \frac{1}{2}]_h,$$

$$\widehat{u_h}(\omega) = \mathbf{FT}(A_h P_q^0 u_H)(\omega) = \widehat{A_h}(\omega)\mathbf{FT}(P_q^0 u_H)(\omega)$$

$$= q^{-1}\widehat{A_h}(\omega)\widehat{u_H}(\omega) = \frac{1}{2}(1+\cos(\omega h))\widehat{u_H}(\omega) = cos^2(\omega h/2)\widehat{u_H}(\omega), \quad \omega \in T_h^1 .$$

(Notice that for $\widehat{u_H}$, originally defined on T_H^1, now its periodic continuation on T_h^1 is used.) We can write this also as

$$\vec{\mathbf{FT}}(u_h) = \begin{pmatrix} \cos^2(\omega h/2) \\ \sin^2(\omega h/2) \end{pmatrix} \widehat{u_H}\,, \quad \omega \in T_H^1.$$

Analogous to (7.31) we introduce the notation

$$\vec{\mathbf{FT}}(u_h)(\omega) = \mathbf{FT}(P_q(a_h))(\omega) \;\; \mathbf{FT}(u_H)(\omega)\,. \tag{7.35}$$

Here $\vec{\mathbf{FT}}(P_q(a_h))$ is a $Q \times 1$- matrix, given by

$$\vec{\mathbf{FT}}(u_h)(\omega) = \mathbf{FT}(P_q(a_h)(\omega) = \frac{1}{Q}\begin{pmatrix} \widehat{A_h}(\omega) \\ \vdots \\ \widehat{A_h}(\omega + 2\pi/q_h) \\ \vdots \end{pmatrix}, \quad \omega \in T_H^n,$$

and $\vec{\mathbf{FT}}(u_h)(\omega)$ a Q-vector.

7.9 The order of a restriction or a prolongation

In the previous section we saw that a restriction or a prolongation could be associated with a gridfunction a_h, an operator A_h or its FT $\widehat{A_h}(\omega)$. Such a FT is a trigonometric polynomial in $\theta = \omega h$. If there is no possibility of confusion in this section we write $A(\theta)$ instead of $\widehat{A_h}(\omega)$ (for restrictions) or $\frac{1}{Q}\widehat{A_h}(\omega)$ for prolongations.

Definition 7.24 The *(primary) order* or *low-frequency order* of a prolongation $P_q(a_h)$ or of a restriction $R_q(a_h)$, is the largest number $m \geq 0$ for which

$$A(\theta) = 1 + \mathcal{O}(|\theta|^m), \text{ for } |\theta| \to 0\,.$$

Definition 7.25 For a given grid-coarsening factor $q \in \mathbb{Z}^n$, $H = qh$, the *secondary order* or *high-frequency order* of a prolongation $P_q(a_h)$ or of a restriction $R_q(a_h)$, is the largest number $m \geq 0$ for which

$$A(\theta + 2\pi p/q) = \mathcal{O}(|\theta|^m)\;(|\theta| \to 0),$$

for all $p \in [0, q)^n \subset \mathbb{Z}^n, p \neq 0^n.$

The use of these two definitions will become clear in section 7.11 where primary and secondary orders play a role in the convergence of multigrid methods. We will now give a few examples of the computation of primary and secondary orders of some prolongations and restrictions. In all examples we assume $q = 2$.

Example 7.26 *Canonical q-restriction* $R_q^0, n = 1$

$$A_h = a_h\star = [0,1,0]_h,$$

$A(\theta) = 1$, hence the primary order is ∞ and the secondary order is 0.

Example 7.27 *Linear interpolation,* $n = 1$

$$A_h = a_h\star = [1/2, 1, 1/2]_h,$$

$$\frac{1}{Q}\widehat{A_h}(\omega) = \cos^2(\omega h/2) \quad \text{(see example 7.23)},$$

$$A(\theta) = \cos^2(\theta/2) = 1 - \frac{1}{4}\theta^2 + \mathcal{O}(\theta^4) \quad \text{for} \quad \theta \to 0\,,$$

hence the primary order is also 2.

$$A(\theta+\pi) = \cos^2(\frac{\theta+\pi}{2}) = \sin^2(\theta/2) = \frac{1}{4}\theta^2 + \mathcal{O}(\theta^4) \quad \text{for} \quad \theta \to 0,$$

hence the secondary order is also 2.

Example 7.28 *Linear interpolation,* $n = 2$

$$A_h = a_h\star = \begin{bmatrix} 1/2 & 1/2 & \\ 1/2 & 1 & 1/2 \\ & 1/2 & 1/2 \end{bmatrix}_h .$$

We write $h\omega = (\phi, \theta)$ and find

$$\begin{aligned} 4A(\phi,\theta) = \;& 1 + \tfrac{1}{2}e^{i\phi} + \tfrac{1}{2}e^{-i\phi} + \tfrac{1}{2}e^{i\theta} + \tfrac{1}{2}e^{-i\theta} + + \tfrac{1}{2}e^{i(\phi-\theta)} + \tfrac{1}{2}e^{-i(\phi-\theta)} \\ &= 1 + \cos\phi + \cos\theta + \cos(\phi-\theta) \\ &= (1+\cos\theta)(1+\cos\phi) + \sin\phi\sin\theta \\ &= 2\cos^2(\theta/2)2\cos^2(\phi/2) + 2\sin(\phi/2)\cos(\phi/2)\,2\sin(\theta/2)\cos(\theta/2) \\ &= 4\cos(\theta/2)\cos(\phi/2)\cdot[\cos(\phi/2)\cos(\theta/2) + \sin(\phi/2)\sin(\theta/2)] \\ &= 4\cos(\theta/2)\cos(\phi/2)\cos((\phi-\theta)/2). \end{aligned}$$

Expanding $A(\phi,\theta)$ in a Taylor series yields

$$\begin{aligned} A(\phi,\theta) \;&= (1 - \tfrac{1}{2}(\tfrac{\theta}{2})^2 + \mathcal{O}(|\theta|^4))\cdot(1 - \tfrac{1}{2}(\tfrac{\phi}{2})^2 + \mathcal{O}(|\phi|^4))\cdot(1 - \tfrac{1}{2}(\tfrac{\phi-\theta}{2})^2 + \mathcal{O}(|\phi-\theta|^4)) \\ &= 1 - \tfrac{1}{4}(\theta^2 + \phi^2 - \phi\theta) + \mathcal{O}((\phi,\theta)^4), \quad \text{for} \quad (\phi,\theta) \to 0\,. \end{aligned}$$

We conclude that the primary order of linear interpolation in two dimensions is 2. For the higher harmonics we find

$$A(\pi-\phi,\theta) = \frac{1}{4}(\theta\phi - \phi^2) + \mathcal{O}((\phi,\theta)^4), \quad \text{for} \;\; (\phi,\theta) \to 0,$$

and similar expressions for $A(\phi, \pi-\theta)$ and $A(\pi-\phi, \pi-\theta)$. We conclude that the secondary order is 2.

Remark: The same holds for the linear weighted restriction (7-point restriction).

Example 7.29 *Bilinear interpolation, $n = 2$*

$$A_h = a_h\star = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}_h .$$

$$\begin{aligned} A(\phi,\theta) &= 1 + \tfrac{1}{2}e^{i\phi} + \tfrac{1}{2}e^{-i\phi} + \tfrac{1}{2}e^{i\theta} + \tfrac{1}{2}e^{-i\theta} + \\ + &\quad \tfrac{1}{4}e^{i(\phi-\theta)} + \tfrac{1}{4}e^{-i(\phi-\theta)} + \tfrac{1}{4}e^{i(\phi+\theta)} + \tfrac{1}{4}e^{-i(\phi+\theta)} \\ &= 1 + \cos\phi + \cos\theta + \tfrac{1}{2}\cos(\phi-\theta) + \tfrac{1}{2}\cos(\phi+\theta) \\ &= 1 + \cos\phi + \cos\theta + cos\phi\cos\theta \\ &= (1+\cos\phi)\,(1+\cos\theta) \\ &= 4\cos^2(\phi/2)\cos^2(\theta/2)\,. \end{aligned}$$

Expanding $A(\phi,\theta)$ in a Taylor series we find

$$A(\phi,\theta) = (1 - (\frac{\phi}{2})^2 + \mathcal{O}(|\phi|^4))(1 - (\frac{\theta}{2})^2 + \mathcal{O}(|\theta|^4))$$

$$= 1 - \frac{1}{4}(\phi^2 + \theta^2) + \mathcal{O}((\phi,\theta)^4) \quad \text{for} \quad (\phi,\theta) \to 0\,.$$

The primary order of bilinear interpolation is 2. Similarly we find that the secondary order is also 2.

Remark: The same holds for the full-weighting (FW) restriction (9-point restriction).

Example 7.30 *Cubic interpolation, $n = 1$*

$$A_h = a_h\star = \left[-\frac{1}{16}, 0, \frac{9}{16}, 1, \frac{9}{16}, 0, -\frac{1}{16}\right]_h .$$

$$\begin{aligned} A(\phi) &= \tfrac{1}{Q}\widehat{A_h}(\phi/h) = \tfrac{1}{2}\widehat{A_h}(\phi/h) \\ &= \tfrac{1}{16}[8 + 9\cos\phi - \cos 3\phi] \\ &= \tfrac{1}{16}[8 + 9\cos\phi - (4\cos^3\phi - 3\cos\phi)] \\ &= \tfrac{1}{16}[8(1+\cos\phi) + 4\cos\phi(1-\cos^2\phi)] \\ &= \tfrac{1}{16}[16\cos^2(\phi/2) + 4\cos\phi\sin^2\phi] \\ &= \tfrac{1}{16}[16\cos^2(\phi/2) + 16\cos\phi\sin^2(\phi/2)\cos^2(\phi/2)] \\ &= \cos^2(\phi/2)[1 + \cos\phi\sin^2(\phi/2)] \\ &= (1 - \tfrac{1}{2}(\phi/2)^2 + ...)^2[1 + (1 + ...)((\phi/2)^2 + ...)] \\ &= 1 + \mathcal{O}(|\phi|^4) \quad \text{for} \quad \phi \to 0\,. \end{aligned}$$

The primary order of cubic interpolation in 1-D is 4. Similar considerations for $A(\pi - \phi)$ yield secondary order 4.

Remark: One can show that also for cubic interpolation in two dimensions, both primary and secondary order are 4.

Example 7.31 *Half-weighting, $n = 2$*

$$A_h = a_h\star = \begin{bmatrix} & 1/8 & \\ 1/8 & 1/2 & 1/8 \\ & 1/8 & \end{bmatrix}_h .$$

$$\begin{array}{ll}
\widehat{A_h}(\phi/h,\theta/h) & = \frac{1}{2} + \frac{1}{8}e^{i\phi} + \frac{1}{8}e^{-i\phi} + \frac{1}{8}e^{i\theta} + \frac{1}{8}e^{-i\theta} \\
& = \frac{1}{2} + \frac{1}{4}\cos\phi + \frac{1}{4}\cos\theta \\
A(\phi,\theta) & = \frac{1}{Q}\widehat{A_h}(\phi/h,\theta/h) \\
& = \frac{1}{4}\widehat{A_h}(\phi/h,\theta/h) \\
& = \frac{1}{16}[2 + cos\phi + \cos\theta] \\
& = \frac{1}{8}(\cos^2(\phi/2) + \cos^2(\theta/2)) \\
& = \frac{1}{8}[1 - \frac{1}{2}(\phi/2)^2 + 1 - \frac{1}{2}(\theta/2)^2 + \mathcal{O}((\phi,\theta)^4)] \\
& = \frac{1}{4} - \frac{1}{64}(\phi^2+\theta^2) + \mathcal{O}((\phi,\theta)^4) \quad \text{for} \quad (\phi,\theta)\to 0,
\end{array}$$

hence the primary order is 2.

$$\begin{array}{ll}
A(\pi-\phi,\theta) & = \frac{1}{8}(\sin^2(\phi/2) + \cos^2(\theta/2)) \\
& = \frac{1}{8}(\frac{1}{2}(\phi/2)^2 + 1 - \frac{1}{2})^2 + \mathcal{O}((\phi,\theta)^4) \\
& = \frac{1}{8} + \frac{1}{64}(\phi^2-\theta^2) + \mathcal{O}((\phi,\theta)^4) \quad \text{for} \quad (\phi,\theta)\to 0,
\end{array}$$

hence the secondary order is 0.

7.10 Fourier analysis in the case of mutual influencing frequencies

In Section 7.8 we saw that $\widehat{u_h}(\omega)$, $\omega\in T_h^n$ could also be denoted by a Q-vector $\vec{\mathbf{FT}}(u_h)(\omega)$, $\omega\in T_H^n$ by taking the harmonic frequencies together

$$\vec{\mathbf{FT}}(u_h)(\omega) = \begin{pmatrix} \widehat{u_h}(\omega) \\ \vdots \\ \widehat{u_h}(\omega + 2\pi p/qh) \\ \vdots \end{pmatrix}, \quad \omega\in T_H^n = T_{qh}^n,\ p\in[0,q). \tag{7.36}$$

Consistent with this notation we also introduce the notation

$$\vec{\mathbf{FT}}(A_h)(\omega) = \text{ diag } (\widehat{A_h}(\omega + 2\pi p/qh)), \quad \omega\in T_{qh}^n,\ p\in[0,q) \tag{7.37}$$

to replace $\widehat{A_h}(\omega)$, $\omega\in T_h^n$.
Remark: $\widehat{A_h}(\omega)\widehat{u_h}(\omega)$, $\omega\in T_h^n$, is now denoted as $\vec{\mathbf{FT}}(A_h)(\omega)\ \vec{\mathbf{FT}}(u_h)(\omega)$, $\omega\in T_H^n$.

The above notation enables us to apply the usual Fourier analysis in the case of mutual influencing of harmonic frequencies. For example, let A_H be a Galerkin-approximation of A_h, i.e.

$$A_H = R_{Hh}A_hP_{hH}, \tag{7.38}$$

where R_{Hh} and P_{hH} are homogeneous restrictions and prolongations, and A_h a Toeplitz operator. We find

$$\mathbf{FT}(A_H)(\omega) = \vec{\mathbf{FT}}(R_{Hh})(\omega)\ \vec{\mathbf{FT}}(A_h)(\omega)\ \vec{\mathbf{FT}}(P_{hH})(\omega), \tag{7.39}$$

where $\vec{\mathbf{FT}}(R_{Hh})(\omega)$ is the $(1\times Q)$-matrix and $\vec{\mathbf{FT}}(P_{hH})(\omega)$ the $(Q\times 1)$- matrix (cf. (7.31) and (7.35)).

Similarly we find for the coarse-grid-correction (CGC) operator

$$M_h^{CGC} = I_h - P_{hH}L_H^{-1}R_{Hh}L_h \tag{7.40}$$

the following FT:

$$\vec{\mathbf{FT}}(M_h^{CGC})(\omega) = I - \vec{\mathbf{FT}}(P_{hH})(\omega)\ \mathbf{FT}(L_H^{-1})(\omega)\ \vec{\mathbf{FT}}(R_{Hh})(\omega)\ \vec{\mathbf{FT}}(L_h)(\omega),\ \omega \in T_H^n\,. \tag{7.41}$$

For $n = 2, q = 2$ (hence $Q = 4$) (7.41) is a (4×4)- matrix:

Because $\|\widehat{u_h}\|_{L^2(T_h^n)} = \|u_h\|_{l_h^2(\mathbb{Z}_h^n)}$, we have the following relation for the norms of the amplification operators:

$$\begin{aligned}\|M_h\|_{2\leftarrow 2} &= \sup_{u_h \in l_h^2(\mathbb{Z}_h^n)} \frac{\|M_h u_h\|_2}{\|u_h\|_2} = \\ &= \sup_{u_h} \frac{\|\ \vec{\mathbf{FT}}(M_h u_h)\|_{L^2}}{\|\ \vec{\mathbf{FT}}(u_h)\|_{L^2}} = \sup_{\widehat{u_h}} \frac{\|\ \vec{\mathbf{FT}}(M_h)(\omega)\ \widehat{u_h}(\omega)\|_{L^2(T_h^n)}}{\|\widehat{u_h}(\omega)\|_{L^2(T_h^n)}} \\ &\le \sup_{\omega \in T_H^n} \|\ \vec{\mathbf{FT}}(M_h)(\omega)\|_2\,.\end{aligned}$$

Since there is no coupling between non-harmonic-related frequencies we also have

$$\rho(M_h) = \sup_{\omega \in T_H^n} |\lambda_{\max}(\ \vec{\mathbf{FT}}(M_h)(\omega))|\,. \tag{7.42}$$

7.11 Requirements for transfer operators

Let L be an ordinary differential operator of order M:

$$L = \sum_{m=0}^{M} c_m \frac{d^m}{dx^m}. \tag{7.43}$$

Its FT or symbol is then given by

$$\widehat{L}(\omega) = \sum_{m=0}^{M} c_m (i\omega)^m, \tag{7.44}$$

(Section 7.6). Let L_h and L_H be two discretisations of L of order $\tilde{p}$ on $\mathbb{Z}_h^n$ and $\mathbb{Z}_H^n$ respectively. Then $\widehat{L_h}(\omega)$ is a trigonometric polynomial of degree M in ω:

$$\widehat{L_h}(\omega) = \sum_{m=0}^{M} \frac{c_m}{h^m} S_m(h\omega), \tag{7.45}$$

where S_m is the FT of the m^{th} order part of the discrete operator L_h. Furthermore, we have

$$|\widehat{L_h}(\omega) - \widehat{L}(\omega)| = \mathcal{O}(h^{\tilde{p}}) \text{ for } h \to 0\,, \tag{7.46}$$

(see (7.25)) and

$$\begin{aligned} S_m(\omega h) &\to (\omega h)^m, \quad \text{for} \quad \omega h \to 0, \\ S_m(\omega h) &= \mathcal{O}(1), \quad (-\pi \le \omega h \le \pi). \end{aligned}$$

We consider a fixed ω and $h \to 0$ and we define $\theta = \omega h$.

Using the above we find:

$$\left\|\frac{\widehat{L_h}(\omega)}{\widehat{L_H}(\omega)}\right\| = \left\|\frac{\widehat{L}(\omega) + \mathcal{O}(h^{\tilde{p}})}{\widehat{L}(\omega) + \mathcal{O}(H^{\tilde{p}})}\right\| = 1 + \mathcal{O}(h^{\tilde{p}})(h \to 0),$$

and
$$\left\| \frac{\widehat{L_h}((\omega + 2\pi p/qh)}{\sum_{m=0}^{M} \frac{c_m}{(qh)^m}(\omega qh)^m} \right\| = \frac{\mathcal{O}(h^{-M})}{\mathcal{O}(h^{-M}(\omega h)^M)} = \mathcal{O}(h^{-M}) \text{ for } h \to 0) .$$

Now we are interested in the behaviour of the CGC-amplification factors for harmonic frequencies in a neighbourhood of the origin. For simplicity we restrict ourselves to the case $n = 1, q = 2$. The amplification operator for the error in one CGC-step is given by (7.40), its FT, $\widehat{M_h^{CGC}}(\omega)$, is given in (7.41) and can now be denoted as

$$\widehat{M}(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \widehat{P}(\omega) \\ \widehat{P}(\omega + \pi/h) \end{pmatrix} \widehat{L_H^{-1}}(\omega)(\widehat{R}(\omega), \widehat{R}(\omega{+}\pi/h)) \begin{pmatrix} \widehat{L_h}(\omega) & 0 \\ 0 & \widehat{L_h}(\omega + \pi/h) \end{pmatrix} .$$

If we assume that the primary and secondary orders of the prolongation (restriction) are given by m_1 and m_2 (n_1 and n_2), then we find for $\theta \to 0$:

$$\begin{aligned} \widehat{M}(\theta) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 + \mathcal{O}(\theta^{m_1}) \\ \mathcal{O}(\theta^{m_2}) \end{pmatrix} (1 + \mathcal{O}(\theta^{n_1}), \mathcal{O}(\theta^{n_2})) \begin{pmatrix} 1 + \mathcal{O}(\theta^{\hat{p}}) & 0 \\ 0 & \mathcal{O}(\theta^{-M}) \end{pmatrix} \\ &= \begin{pmatrix} \mathcal{O}(\theta^{m_1}) + \mathcal{O}(\theta^{n_1}) + \mathcal{O}(\theta^{\hat{p}}) & \mathcal{O}(\theta^{n_2 - M}) \\ \mathcal{O}(\theta^{m_2}) & 1 + \mathcal{O}(\theta^{n_2 + m_2 - M}) \end{pmatrix} . \end{aligned}$$

It is clear that convergence requires $m_1 > 0, n_1 > 0$ and $\tilde{p} > 0$. Further, for $\theta \to 0$ the eigenvalues of $M(\theta)$ are

$$\lambda_1 = 1 + \ldots , \text{ and } \lambda_2 = \mathcal{O}(\theta^{n_2 + m_2 - M}).$$

For convergence of the CGC-process $n_2 + m_2 \geq M$ is required. Further, $n_2 \geq M$, $m_2 \geq 0$ is required for $\|M(\theta)\|$ to be bounded. Similarly we derive the requirements $m_2 \geq M$, $n_2 \geq 0$ for $\|\overline{M}(\theta)\|$ to be bounded.