Semi-algebraic multi-level methods based on wavelet decompositions I: Application to two-point boundary value problems

Andreas Rieder *
Rice University
Department of Mathematics
Houston, TX 77251

April 1993

Abstract

To clarify more precisely the vague but often indicated connection between wavelet and multi-grid theory is the main goal of the article. Therefore we present a multi-level method based on a wavelet approximation of the succesive error of a classical iterative solver. The resulting iteration is a hybrid between a purely algebraic multi-level technique and the usual multi-grid technique related to a discretization of an elliptic differential operator. It shows the typical multi-grid convergence behaviour, even if it is not able to compete at this stage with the efficient multi-grid iterations. Nevertheless, our new approach has the capacity to solve linear equations arising from the discretization of integral operators of the first kind by multi-level techniques.

Key words: wavelet, multi-level iteration, multi-grid technique

Subject classification: AMS(MOS) 65F10

*supported by a Feodor Lynen-Fellowship of the Alexander von Humboldt-Foundation

1 Introduction

Wavelet theory is a new and fast developing field of mathematical research. Besides being used in mathematics itself (e.g. unconditional bases of certain function spaces [16], analysis of fractal objects [14]) wavelets are now used in a wide range of applications. We mention only a few: They led to fast analysis and reconstruction algorithms to reduce the storage costs of visual data [8]. In signal analysis they are well suited for a phase space representation of signals [6] as well as the detection of edges and singularities [10,19]. Because of their localization in phase space they are also of interest in quantum field theory [1].
Meanwhile - in the last three years - wavelets have appeared in numerical analysis. But their applications in this field are mostly theoretical so far. Some others are practical, e.g. in [2] the authors introduce a rapid algorithm for the application of a class of linear operators to arbitrary vectors. They take advantage of the same property which enables the reduction of visual data. There are also serious attempts to solve PDEs by wavelet techniques [26,27]. Indeed, the numerical solution of PDEs will be treated more and more by these methods. In particular the special structure of wavelets raises hopes for obtaining algorithms for the adaptive grid refinement.

In this paper we study the connection between (discrete) wavelets and multi-level resp. multi-grid methods. Everyone who is little familar with both fields, the wavelet as well as the multi-level theory, feels intuitively the strong correspondence between them. The recent article [3] states some general reasons for that. The authors base their discussion on the point that the linear system to be solved arises from a Wavelet-Galerkin discretization of an elliptic PDE which, for instance, is the case in [26] leading to a wavelet-multi-grid solver [9]. To avoid the many problems which may be caused by this approach (e.g. the modeling of the boundary data) we chose the more convenient algebraic one. And we were proved right: Our wavelet based prolongation and restriction operators lead to multi-level iterations with the typical dimension (discretization step size) independent convergence rate. Nevertheless, the introduced algorithms are at this stage *not* a real alternative to the still more efficient multi-grid solvers. That was not our goal. Indeed, the intention of this paper was to clarify more precisely the above mentioned intuition. However we are convinced that a modification of our wavelet multi-level algorithm will be able to compete with the multi-grid derivations [12] for the treatment of anisotropic (singularly perturbed) 2D or 3D problems. First numerical experiments are encouraging.

Our proposed method works because wavelet analysis separates the low frequency (smooth or global) from the high frequency (rough or local) parts of a vector. Both resulting vectors have about half the dimension of the original one. Again the vector containing the low frequency parts can be decomposed in the same way. Applying the procedure recursively gives a lower dimensional vector which represents the smooth core of the starting vector and a sequence of vectors with increasing dimension which contain all the details of a certain order of magnitude. The original vector can be synthesized by these lower dimensional vectors without loss of information. This separation into low and high frequency components predestines the wavelet operators to serve as a fine-to-coarse and a coarse-to-fine transfer. The classical iterative solvers for linear equations show the interesting behaviour: their successive errors are nearly linear combinations of eigenvectors belonging to small eigenvalues of the underlying matrix. Consequently, for some matrices the errors are smooth, for others the errors are high frequency in the wavelet sense and so, they are determined by lower dimensional vectors. If we carry out the usual multi-grid procedure [11] with the 'wavelet prolongation' and 'wavelet restriction', then we get a multi-level method with the same structure as the (continuous) multi-grid methods. In contrast to the purely algebraic multi-grid iterations [24], which only need information contained completely in the matrix, the right choice of the prolongation resp. restriction operators in our method depends on the origin of the linear equation. Of course we are also free of any geometric restriction in the definition of the fine-to-coarse and the coarse-to-fine transfer. Therefore we call our method a *semi-algebraic multi-level method*.

In the next section we prepare the infinite dimensional wavelet operators of Daubechies [5] for our finite dimensional purpose, and we list some important properties of these 'new' operators. Section 3 deals with the general two-level method, introduces the wavelet two-level method and considers it theoretically with respect to its application to two-point boundary problems. The step to the multi-level method is straightforward and done in Section 4. Finally, in this section numerical results are presented for the comparision of the wavelet multi-level with the usual multi-grid method.

2 Finite dimensional wavelet operators

2.1 Short introduction to discrete wavelets

In her paper '*Orthonormal Bases of Compactly Supported Wavelets*' [5], see also [7], Daubechies constructs continuous linear operators H_N, G_N : $l^2(\mathbb{Z}) \to l^2(\mathbb{Z})$, $N \in \mathbb{N}$, with useful properties:

$$
\begin{aligned}
H_N^* H_N \,+\, G_N^* G_N &= I_{l^2}\,, & (1)\\
H_N H_N^* \,=\, G_N G_N^* &= I_{l^2}\,, & (2)\\
G_N H_N^* \,=\, H_N G_N^* &= 0\,. & (3)
\end{aligned}
$$

I always represents the identity operator. The third equation is equivalent to

$$\overline{\text{range } H_N^*} \perp \overline{\text{range } G_N^*}\,.$$

Each of the operators H_N and G_N is uniquely determined by filter functions h_N, $g_N : \mathbb{Z} \to \mathbb{R}$ via

$$
\begin{aligned}
(H_N x)_k &= \sum_{n\in\mathbb{Z}} h_N(n-2k)\, x_n\,,\\
(G_N x)_k &= \sum_{n\in\mathbb{Z}} g_N(n-2k)\, x_n\,,
\end{aligned}
\qquad (4)
$$

where h_N and g_N fulfill:

$$h_N(l) \,=\, 0 \quad \text{for} \quad l \notin \{0,\ldots,2N-1\}\,, \qquad g_N(l) \,=\, (-1)^l h_N(1-l)\,,$$

$$\sum_{l=0}^{2N-1} h_N(l) \,=\, \sqrt{2}\,, \qquad \sum_{l=2-2N}^{1} l^k g_N(l) \,=\, 0 \quad \text{for} \quad k \in \{0,\ldots,N-1\}\,.$$

So, H_N may be interpreted as an averaging or smoothing operator whose smoothing effect increases with N. Conversely, G_N can be viewed (up to a factor) as a difference approximation to the ordinary differential operator D^N of order N and due to the vanishing moments property of the filter g_N the approximation G_N is exact for polynomials of maximal degree $N-1$.
Any sequence $c^0 \in l^2(\mathbb{Z})$ can be decomposed into two sequences

$$
\begin{aligned}
c^1 &= H_N c^0\,, & (5)\\
q^1 &= G_N c^0\,. & (6)
\end{aligned}
$$

The sequences in (5),(6) contain 'half the information' of c^0 because of the scaling factor 2 in (4). Further they are not redundant because of $c^0 = H_N^* c^1 + G_N^* q^1$ and $\langle H_N^* c^1, G_N^* q^1 \rangle_{l^2} = 0$. c^1 is a smoothed version of c^0 and q^1 represents the high frequency or rough parts. The *wavelet decomposition* of c^0 is defined recursively by

$$\left.\begin{array}{rcl} c^{i+1} &=& H_N c^i \,, \\ q^{i+1} &=& G_N c^i \,. \end{array}\right. \qquad i = 0, 1, \ldots \tag{7}$$

If we stop the analysis at *level m* we can synthesize c^0 by c^m, q^m, q^{m-1},..., q^1 via

$$c^{i-1} = H_N^* c^i + G_N^* q^i \,, \quad i = m, \ldots, 1 \,. \tag{8}$$

The difference between c^0 and the highly smoothed c^m is hidden in the coefficients q^j which contain all the information lost by the transition from c^{i-1} to its smoothed version c^i. Again c^{i-1} splits into its global part c^i and its local part (wavelet-detail) q^j.
Surely our short introduction in the wavelet field leaves many questions unanswered, especially the continuous background (which explains the name 'wavelet') and the great advantage of the wavelet analysis for digital signal analysis , image encoding, etc. For these we refer to [5,13,16,23]. However we have mentioned all properties and interpretations of H_N and G_N which are essential for our purpose. In the rest of this paragraph we modify H_N and G_N to become linear mappings from $\mathbb{R}^n$ into $\mathbb{R}^m$, $m \approx n/2$, and motivate the possibility to use them as a restriction operator for the fine-to-coarse transfer resp. as a prolongation operator for the coarse-to-fine transfer.

2.2 Construction of the finite dimensional wavelet operators

We introduce the canonical embedding $P_n : \mathbb{R}^n \to l^2(\mathbb{Z})$ by

$$(P_n x)_k = \left\{ \begin{array}{ccc} x_k & : & 0 \le k \le n-1 \\ 0 & : & \text{otherwise} \end{array} \right. ,$$

and the shift operator $S_t : l^2(\mathbb{Z}) \to l^2(\mathbb{Z})$, $t \in \mathbb{Z}$, by

$$(S_t x)_k = x_{k-t} \,.$$

For our further calculations we need some obvious identities,

$$P_n^* P_n \;=\; I_{I\!R^n}\,, \qquad P_n P_n^* \;=\; \mathcal{P}_n\,,$$

$$S_{-t} S_t \;=\; S_t^* S_t \;=\; I_{l^2}\,,$$

where $\mathcal{P}_n$ denotes the orthogonal projection onto the closed subspace range(P_n). We denote the 'greatest integer' by bracket symbols $[\cdot]$. Throughout this paper we start the numbering of vector and matrix entries at 0.

Lemma 2.1 *The finite dimensional Daubechies operators,*

$$H_N^n \;:=\; P^*_{N+\left[\frac{n-1}{2}\right]} S_{N-1}\, H_N\, P_n \,:\, I\!R^n \to I\!R^{N+\left[\frac{n-1}{2}\right]}\,, \tag{9}$$

$$G_N^n \;:=\; P^*_{N+\left[\frac{n-1}{2}\right]}\, G_N\, P_n \,:\, I\!R^n \to I\!R^{N+\left[\frac{n-1}{2}\right]}\,, \tag{10}$$

satisfy:

a) $\quad (H_N^n)^*\, H_N^n \;+\; (G_N^n)^*\, G_N^n \;=\; I_{I\!R^n}\,,$

b) $\quad (H_N^n)^*$ *and* $(G_N^n)^*$ *are injective.*

Remark: Unfortunately, the properties (2) and (3) of H_N and G_N are not transmitted to H_N^n and G_N^n. That is because the fact that the range of P_n is not invariant under the action of $H_N^* H_N$ resp. $G_N^* G_N$.

Proof: a) First we verify the implication

$$y \in \mathrm{range}(H_N P_n) \quad\Longrightarrow\quad y_k = 0 \quad \text{if} \quad k \le -N \quad \text{or} \quad k \ge \left[\tfrac{n-1}{2}\right] + 1\,. \tag{11}$$

We have

$$y_k = (H_N P_n x)_k = \sum_{l=0}^{2N-1} h_N(l)\,(P_n x)_{2k+l}$$

with

$$(P_n x)_{2k+l} = \begin{cases} x_{2k+l} & : \quad 0 \le 2k+l \le n-1 \\ 0 & : \quad \text{otherwise} \end{cases}\,.$$

Here $k \le -N$ yields $2k+l \le -2N+2N-1 = -1$ for all $0 \le l \le 2N-1$ and hence $(H_N P_n x)_k = 0$ since $(P_n x)_{2k+l} = 0$ for $0 \le l \le 2N-1$. Analogously we get $(H_N P_n x)_k = 0$ for $k \geq \left[\frac{n-1}{2}\right] + 1$. As an obvious consequence of (11) we observe that

$$\text{range}\,(S_{N-1} H_N P_n) \subset \text{range}\left(P_{N+[\frac{n-1}{2}]}\right) .$$

In the same way we find

$$\text{range}\,(G_N P_n) \subset \text{range}\left(P_{N+[\frac{n-1}{2}]}\right) ,$$

and finally obtain

$$\begin{aligned}(H_N^n)^* H_N^n + (G_N^n)^* G_N^n &= P_n^* H_N^* S_{1-N} \underbrace{P_{N+[\frac{n-1}{2}]} P^*_{N+[\frac{n-1}{2}]}}_{=\mathcal{P}_{N+[\frac{n-1}{2}]}} S_{N-1} H \\ &\quad + P_n^* G_N^* \underbrace{P_{N+[\frac{n-1}{2}]} P^*_{N+[\frac{n-1}{2}]}}_{=\mathcal{P}_{N+[\frac{n-1}{2}]}} G_N P_n \\ &= P_n^* (H_N^* H_N + G_N^* G_N) P_n = P_n^* P_n = I_{\mathbb{R}^n} .\end{aligned}$$

b) If we rewrite (2) to become

$$\sum_{k \in \mathbb{Z}} g_N(k-2l)\, g_N(k-2m) = \sum_{k \in \mathbb{Z}} h_N(k-2l)\, h_N(k-2m) = \delta_{l\,m}$$

we see that the columns of the matrices $(H_N^n)^*$ and $(G_N^n)^*$ are linearly independent. ♣

Any element $v = v^0 \in \mathbb{R}^n$ can be decomposed into

$$\begin{aligned} v^1 &= H_N^{n_0}\, v^0 \in \mathbb{R}^{n_1} , \\ q^1 &= G_N^{n_0}\, v^0 \in \mathbb{R}^{n_1} \end{aligned}$$

with $n_0 = n,\ n_1 = N + \left[\frac{n_0 - 1}{2}\right]$ and

$$v^0 = (H_N^{n_0})^*\, v^1 + (G_N^{n_0})^*\, q^1 .$$

According to the remark after Lemma 2.1 and in contrast to (5) and (6) the above splitting of the vector v into two lower dimensional vectors is redundant. If $n_1 > N + \left[\frac{n_1 - 1}{2}\right] =: n_2$ we can go on to analyze v^1 by

$$\begin{aligned} v^2 &= H_N^{n_1}\, v^1 \in \mathbb{R}^{n_2} , \\ q^2 &= G_N^{n_1}\, v^1 \in \mathbb{R}^{n_2} , \end{aligned}$$

and $v = v^0$ satisfies

$$v = (H_N^{n_0})^*(H_N^{n_1})^* \, v^2 \; + \; (H_N^{n_0})^*(G_N^{n_1})^* \, q^2 \; + \; (G_N^{n_0})^* \, q^1 \, .$$

We terminate the decomposition of v into lower dimensional vectors if we reach n_l with $n_l \; = \; N \; + \; \left[\frac{n_l - 1}{2}\right]$: the operators $H_N^{n_l}$ and $G_N^{n_l}$ are isomorphisms for this special dimension. An explicit expression of $\{n_l\}_{l \in I\!N_0}$ is stated in

Lemma 2.2 *Let $N \geq 1$. The inductively defined sequence*

$$n_{k+1} \; = \; N + \left[\frac{n_k - 1}{2}\right] \; , \quad n_0 \geq 2N - 1$$

can explicitly expressed by

$$n_k \; = \; 2N - 1 \; + \; \left[\frac{n_0 - 2N + 1}{2^k}\right] \; .$$

Proof: We claim the estimate

$$n_k^{\min} := 2(N-1) + \frac{n_0 - 2(N-1)}{2^k} \; \leq \; n_k \; \leq \; n_k^{\max} := 2N - 1 + \frac{n_0 - 2N + 1}{2^k}$$

which will be verified by induction. For $k = 0$ nothing has to be shown. Assume the estimate is true for some $k \in I\!N$. Then

$$\begin{aligned} n_{k+1} = N + \left[\frac{n_k - 1}{2}\right] \quad & \leq \quad N + \frac{n_k}{2} - \frac{1}{2} \\ & \leq \quad N - \frac{1}{2} + N - \frac{1}{2} + \frac{n_0 - 2N + 1}{2^{k+1}} \\ & = \quad 2N - 1 + \frac{n_0 - 2N + 1}{2^{k+1}} \; = \; n_{k+1}^{\max} \, , \end{aligned}$$

and $\quad n_{k+1} \; \geq \; N + \frac{n_k}{2} - 1 \; \geq \; n_{k+1}^{\min}$.
Since $n_k^{\max} \; - \; n_k^{\min} \; = \; 1 - \frac{1}{2^k} \; < \; 1$, $n_k^{\max}$ and $n_k^{\min}$ enclose at most only one integer and

$$n_k^{\min} \leq \; n_k \; \leq \; [n_k^{\max}] \; \leq \; n_k^{\max}$$

implies $\quad n_k \; = \; [n_k^{\max}]$. ♣

Corollary 2.3 *Suppose $n_0 \; \geq \; 2N - 1$ and $k^* \; = \; \min\{k \in I\!N_0 \; : \; n_0 - 2N + 1 \; < \; 2^k\}$. Then*

$$n_k \; = \; 2N - 1 \quad \textit{for all} \quad k \; \geq \; k^* \, .$$

The lemma below reduces any vector to its global or smooth core and to a hierarchy of details covering that core.

Lemma 2.4 *Any* $v = v^0 \in \mathbb{R}^n$ *with* $n = n_0 \geq 2N-1$ *has the representation*

$$v \;=\; \prod_{k=0}^{l-1} (H_N^{n_k})^* \, v^l \;+\; \sum_{i=1}^{l} \left(\prod_{k=0}^{i-2} (H_N^{n_k})^* \right) (G_N^{n_{i-1}})^* \, q^i \tag{12}$$

where $l = \min\{k \in \mathbb{N}_0 \,:\, n - 2N + 1 \,<\, 2^k\}$*, otherwise the representation of* v *is not defined, and*

$$\begin{aligned} v^l &= H_N^{n_{l-1}} \, H_N^{n_{l-2}} \ldots H_N^{n_0} \, v^0 \quad \in \mathbb{R}^{2N-1}, \\ q^i &= G_N^{n_{i-1}} \, v^{i-1}, \quad v^i \;=\; H_N^{n_{i-1}} \, v^{i-1}, \\ n_i &= N + \left\lceil \frac{n_{i-1} - 1}{2} \right\rceil \;=\; 2N - 1 \;+\; \left\lceil \frac{n - 2N + 1}{2^i} \right\rceil . \end{aligned}$$

Remarks:

a) If n_{i+1} is large relative to N then n_{i+1}/n_i is approximately equal to $1/2$, i.e. the dimension is nearly divided into halves at the transition from the level i to the level $i+1$.

b) In [23] Strang considers only the finite dimensional wavelet operators H_1^n and G_1^n, n even, as an example.

The following interpolation result which is now supplied for later use states more precisely in which sense a vector v is decomposed by (12) into its low frequency part v^l and in parts q^i, $i = l, \ldots, 1$, representing higher and higher frequencies.

Lemma 2.5 *Suppose* u *is an element of the Sobolev space* $H^2(0,1) = \{w \in L^2(0,1) \,|\, w', \, w'' \in L^2\}$ *with* $u(0) = u(1) = 0$ *(*u *is continuous in* $[0,1]$*, cf. Ciarlet [4]). Let the operator* $T_h : H^2(0,1) \to \mathbb{R}^n$ *be defined by* $(T_h u)_k = u((k+1)\,h)$*,* $k = 0, \ldots, n-1$*,* $h = 1/(n+1)$*. If* $N \geq 2$*, then*

$$\| \, (I_{\mathbb{R}^n} \,-\, (H_N^n)^* \, H_N^n) \, T_h u \, \| \;=\; \| \, (G_N^n)^* \, G_N^n \, T_h u \, \| \leq \, c_N \; h \; \|u\|_{H^2} \tag{13}$$

with a positive constant c_N *depending only on* N*.*
$\| \cdot \| = \langle \cdot, \cdot \rangle^{1/2}$ *denotes the usual Euclidean norm of* $\mathbb{R}^n$*.*

Proof: By Taylor's formula

$$u(\cdot + \delta) \;=\; u(\cdot) + \delta\, u(\cdot) + r_\delta\,, \quad |r_\delta| \;\le\; |\delta|^{3/2}\,\|u\|_{H^2}\,,$$

and by the vanishing moments property of the filters g_N we get

$$|(G_N^n\,T_h u)_k| \;\le\; c_{N,1}\;h^{3/2}\,\|u\|_{H^2}\,, \quad \text{for} \quad N-1 \le k \le \left[\frac{n-2}{2}\right]\,,$$

with $c_{N,1} = \sum |g_N(m)|\,|m|^{3/2}$. Next, we consider the indices $0 \le k \le N-2$,

$$|(G_N^n\,T_h u)_k| \le \sum_{l=-2k}^{1} |g_N(l)|\,|u((l+2k+1)\,h)|\,.$$

Using $u(0)=0$ we estimate

$$|u((l+2k+1)\,h)| \;\le\; (l+2k+1)\;h^{1/2}\;\|u'\|_{L^2}\,,$$

which implies

$$|(G_N^n\,T_h u)_k| \;\le\; \underbrace{\max_{0\le k\le N-2}\;\sum_{l=-2k}^{1} |g_N(l)|\,|l+2k+1|}_{=:\,c_{N,2}}\;h^{1/2}\;\|u\|_{H^2}\,.$$

Analogously, we are able to verify

$$|(G_N^n\,T_h u)_k| \;\le\; c_{N,2}\;h^{1/2}\;\|u\|_{H^2}$$

uniformly for $[(n-2)/2]+1 \le k \le n_1 - 1$. The splitting

$$\begin{aligned}
&\|\,G_N^n\,T_h u\,\|^2 \\
&= \sum_{k=0}^{N-2} |\,(G_N^n\,T_h u)_k\,|^2 + \sum_{k=N-1}^{[(n-2)/2]} |\,(G_N^n\,T_h u)_k\,|^2 + \sum_{k=[(n-2)/2]+1}^{n_1-1} |\,(G_N^n\,T_h u)_k\,|^2 \\
&\le \;(2\,c_{N,2}^2\,(N-1) \;+\; c_{N,1}^2)\;h^2\;\|u\|_{H^2}^2
\end{aligned}$$

yields (13). ♣

3 The two-level method

3.1 Introduction and terminology

A general *Galerkin–type two–level method* for the solution of a linear system

$$A\,x \;=\; b\,, \quad A \in I\!R^{n\times n} \text{ positive definite}, \quad b \in I\!R^{n} \tag{14}$$

consists of two components: the *basic iterative method* (BIM) or *smoothing iteration*

$$x^{m+1} \;=\; x^{m} - L^{-1}\,(Ax^{m} - b)\,, \qquad L^{-1} \in I\!R^{n\times n}\,, \tag{15}$$

with *iteration operator* $S \;=\; I - L^{-1}\,A$, and the (Galerkin–type) *coarse grid correction process*

$$y^{\,m+1} \;\;=\;\; y^{m} \;-\; Q^{*}\,(Q\,A\,Q^{*})^{-1}\,Q\,(Ay^{m} - b)\,. \tag{16}$$

The *coarse grid correction operator* $T \;=\; I \;-\; Q^{*}\,(Q\,A\,Q^{*})^{-1}\,Q\,A$ is well defined whenever the *restriction operator* $Q : I\!R^{n} \to I\!R^{k}$, $k \ll n$, has full rank, i.e. the *prolongation operator* Q^{*} is injective. Performing ν_1 steps of (15) before and ν_2 steps after one coarse grid correction step yields the two-level method. The corresponding *two-level iteration operator* reads

$$M^{TL}(\nu_1, \nu_2) \;=\; S^{\nu_2}\,T\,S^{\nu_1}\,. \tag{17}$$

ν_1 and ν_2 are the numbers of *pre-* and *post-smoothing steps*, respectively.
The essential point of the two-level method (TLM) is due to the adequate determination of the restriction resp. the prolongation operator leading to a convergence rate bounded above by a constant smaller than 1 independently of the dimension n.
In case of classical multi-grid methods with a continuous background, where the matrix A is derived by a discretization of an elliptic differential operator, the error $e^{m} \;=\; x^{m} - x^{*}$ of (15) (x^{*} denotes the exact solution of (14)) is smooth since the eigenvectors belonging to small eigenvalues are smooth, in other words, the high oscillating parts of x^{*} are well represented by x^{m}. This *smoothing property* of classical iterative solvers, like the Gauß-Seidel - and the Jacobi iteration etc., is used by (continuous) multi-grid techniques to determine e^{m} by a coarser discretization of the underlying differential operator, Hackbusch [11].
Formally, the algebraic multi-grid approach (AMG) to solve (14) uses the

same principles as the ususal multi-grid methods. However, the coarsening strategy is not based on a priori information of the continuous problem but on information contained completely in the matrix itself. For a comprehensive discussion of AMG we refer to Ruge, Stüben [20]. Further, the authors describe typical problems where the application of standard multi-grid techniques is difficult or imposssible and where the use of AMG becomes very attractive.

For our further considerations we need the *energy-norm* relative to (14)

$$\|x\|_A := \|A^{1/2}x\| \,. \tag{18}$$

The associated matrix norm is $\|B\|_A = \|A^{1/2}BA^{-1/2}\|$. We use the same symbol for vector and their associated matrix norms.
Because the spectral radius of a product does not depend on its order we have

$$\varrho\left(M^{TL}(\nu_1,\nu_2)\right) = \varrho\left(M^{TL}(\nu_1+\nu_2,0)\right) = \varrho\left(M^{TL}(0,\nu_1+\nu_2)\right) \,.$$

Consequently, for a convergence analysis it suffices to study either pre- or post-smoothing. If the BIM converges monotonically with respect to the energy-norm (e.g. Gauß-Seidel -, Jacobi - and Richardson relaxation do) convergence of the TLM is established simply by

$$\varrho\left(M^{TL}(\nu_1,\nu_2)\right) \le \|M^{TL}(\nu_1+\nu_2,0)\|_A \le \|S^{\nu_1+\nu_2}\|_A < 1 \,. \tag{19}$$

The second inequality holds since $A^{1/2}T\,A^{-1/2}$ is an orthogonal projector on $\mathbb{R}^n$ equipped with Euclidean inner product.

It is the aim of this paper to show that the finite dimensional wavelet operators H_N^n resp. G_N^n are good choices for the restriction operator Q in the following sense: the right choice yields a TLM with a contraction number bounded above by a positive constant smaller than 1. Which wavelet operator we will choose depends on the matrix A. The resulting iteration shares features and properties from both, the continuous and the algebraic multi–grid. Therefore it can be viewed as the missing link between them.

3.2 General convergence of two-level methods

We present an abstract convergence theory for the two-level method (17). For convenience we refer to an iterative process by its iteration matrix. The

notation $A < B$ signifies that A and B are symmetric and that $B - A$ is positive definite. The other order relations are defined in the same way.
In the next section the theorem below will be applied to the finite dimensional wavelet operators.

Theorem 3.1 *Let A be the matrix of (14). Suppose the restriction operator $Q : \mathbb{R}^n \to \mathbb{R}^k$, $k < n$, has full rank and satisfies*

$$0 \le I - Q^*Q \le \gamma A \tag{20}$$

with a positive number γ. Further, let the matrix L of (15) be positive definite with $L \ge A$. Then, the contraction number and the spectral radius of the TLM (17) are bounded by

$$\|M^{TL}(\nu_1,\nu_2)\|_A \le \left(1 - \frac{1}{C_A}\right)^{\kappa}, \quad \kappa = \begin{cases} 1/2 & : \quad \nu_1 \ge 1 \;\; or \;\; \nu_2 \ge 1 \\ 1 & : \quad \nu_1,\, \nu_2 \ge 1 \end{cases}, \tag{21}$$

$$\varrho\left(M^{TL}(\nu_1,\nu_2)\right) \le 1 - \frac{1}{C_A} \qquad for\ all \quad \nu_1 + \nu_2 \ge 1 \tag{22}$$

with $C_A = \gamma \|L\|$.

Proof: Note that the first part of (20) implies $\|I - Q^*Q\| \le 1$ as well as $\|(I - Q^*Q)x\|^2 \le \|(I-Q^*Q)^{1/2}x\|^2$. Because of

$$\begin{aligned} \inf_{z \in \mathbb{R}^k} \|L^{1/2}(x - Q^*z)\|^2 &\le \|L^{1/2}(I - Q^*Q)x\|^2 \\ &\le \|L\| \, \langle (I - Q^*Q)x, x \rangle \\ &\overset{(20)}{\le} \gamma \, \|L\| \, \|x\|_A^2 \end{aligned}$$

the theorem follows by Lemma 6.4.8 and Proposition 6.4.10 of [11]. ♣

The theorem above serves our purpose if we use the Richardson, the damped Jacobi or the symmetric Gauß-Seidel iteration as BIM. However, the usual Gauß-Seidel iteration needs an extra consideration.

Lemma 3.2 *We adopt the assumptions on A and Q from the preceeding theorem. Let A be split into $A = D_A - E - F$, $F = E^T$,*

D_A *diagonal matrix (diagonal part of A),*
E *strictly lower triangular matrix.*

If $\|I - D_A^{-1/2} F D_A^{-1/2}\|^2 \le C$*, then the TLM (17) with Gauß-Seidel smoothing satisfies (21) with* $C_A = C\,\gamma\,\|D_A\|$*.*

Proof: The conditions of Proposition 6.4.12 in [11] are satisfied with $C_1 = \sqrt{C}$ and $C_2 = \sqrt{\gamma\,\|D_A\|}$. Hence, the statement of that proposition is just our assertion. ♣

Remark: If the matrix A originates from a discretization of a differential operator, then typically $\|L\|$ as well as $\|D_A\|$ tend to infinity as the discretization step size h decreases. That behaviour will be compensated by $\gamma = \gamma_h$ such that C_A will be bounded uniformly in h, i.e. the bounds for the contraction number (21) and for the spectral radius (22) do not depend on h.

3.3 The wavelet two-level method

The quality of the wavelet two-level method (WTLM) lies in the special choice of the restriction and prolongation operators to define coarse grid corrections

$$y^{m+1} = y^m - (H_N^n)^*\,(H_N^n\,A\,(H_N^n)^*)^{-1}\,H_N^n\,(Ay^m - b)\,, \qquad (23)$$

$$y^{m+1} = y^m - (G_N^n)^*\,(G_N^n\,A\,(G_N^n)^*)^{-1}\,G_N^n\,(Ay^m - b)\,. \qquad (24)$$

To simplify the notation we set $H := H_N^n$, $G := G_N^n$.
Which wavelet coarse grid correction (23) or (24) should be used to solve (14) depends on the matrix $A \in \mathbb{R}^{n\times n}$. Theorem 3.1 and Lemma 3.2 provide the following criterion: If there exists a positive number γ_n fulfilling

$$G^*G \le \gamma_n\,A\,, \qquad (25)$$

then perform the coarse grid correction step (23). Conversely, if

$$H^*H \le \gamma_n\,A \qquad (26)$$

holds, then (24) would be the natural choice for the coarse grid correction.

Remark: The conditions (25) and (26) can be interpreted in terms of the spectral behaviour of A and G^*G resp. H^*H. It is a well known fact that a TLM converges fast, if ('and only if') the BIM and the coarse grid correction reduce error components belonging to complementary parts of the spectrum of A. However, the classical iterative solvers (Richardson, etc.)

damp only those error components corresponding to large eigenvalues pretty well. Clearly the coarse grid correction has to be adapted to the subspace U spanned by the eigenvectors with small eigenvalues. The criterion (25) is satisfied in the case U contains only low frequency vectors. On the other side, in order to hold for (26) the elements in U have to be high frequency. Our criterion tests which wavelet operator reproduces the spectral behaviour of A in the region of small spectral values and so, it guarantees that the wavelet coarse grid correction we choose has the necessary complementary properties to yield an efficient TLM.

3.3.1 WTLM applied to a two-point boundary value problem

We consider the one-dimensional self-adjoint differential equation of second order

$$-\,(q(x)\,u'(x))' \,+\, p(x)\,u(x) \,=\, f(x)\,, \quad x \,\in\, (\alpha,\beta)\,, \tag{27}$$

with boundary conditions

$$u(\alpha) \,=\, u(\beta) \,=\, 0\,. \tag{28}$$

We assume that q and p are given to be continuous in $[\alpha,\beta]$ with

$$0 \,<\, \underline{q} \,\le\, q(x) \,\le \overline{q}\,, \quad 0 \,\le\, p(x) \,\le \overline{p} \quad \text{for} \quad \alpha \,\le\, x \,\le \beta\,. \tag{29}$$

To discretize (27) we use both the Finite-Difference (Finite-Volume) - as well as the Finite-Element (FE) Method with respect to the equidistant grid $\mathcal{G} = \{\, x_i = \alpha + (i+1)\,h \mid -1 \le i \le n,\; h = (\beta-\alpha)/(n+1)\,\}$. The first technique leads to the linear equation of dimension n,

$$A_d\, u_d \,=\, f_d \tag{30}$$

where A_d is the tridiagonal matrix belonging to the 'difference stencil'

$$\frac{1}{h^2}\,[\,-q(x_i^-) \quad q(x_i^-) + q(x_i^+) + h^2\,p(x_i) \quad -\,q(x_i^+)\,]\,, \tag{31}$$

$x_i^- \,=\, x_i - h/2,\; x_i^+ \,=\, x_i + h/2,\; i \,=\, 0,\ldots,n-1$. The right hand side of (30) consists of the components $(f_d)_i = f(x_i)$ and $(u_d)_i$ is an approximation to the solution u of (27) at the point x_i. The matrix A_d is known to be positive definite, cf. Varga [25], and an easy calculation verifies

$$\underline{q}\,h^{-2}\,\mathcal{D} \,\le\, A_d \,\le\, \overline{q}\,h^{-2}\,\mathcal{D} \,+\, \overline{p}\,I \tag{32}$$

where $h^{-2}\,\mathcal{D}$ denotes the matrix of the difference stencil (31) with $q(x) \equiv 1$ and $p(x) \equiv 0$. Indeed, the first inequality of (32) can simply be established by setting $\tilde{q}(x) = q(x) - \underline{q}$, $\tilde{p}(x) = p(x)$ in (27) and the second one by $\tilde{q}(x) = \overline{q} - q(x)$, $\tilde{p}(x) = \overline{p} - p(x)$.
For the Finite-Element approach we denote the standard Sobolev space $H_0^1(\alpha, \beta)$ by

$$H_0^1 \;=\; H_0^1(\alpha,\beta) \;=\; \{\, v \in L^2(\alpha,\beta) \,|\, v' \in L^2 , \quad v(\alpha) = v(\beta) = 0\,\} \,. \qquad (33)$$

The weak (variational) formulation of (27) reads

$$a(u,v) \;=\; (f,v)_{L^2} \qquad \text{for all} \quad v \in H_0^1(\alpha,\beta)\,. \qquad (34)$$

The bilinear form a is defined by

$$a(u,v) \;:=\; \int_\alpha^\beta q(x)u'(x)v'(x) \;+\; p(x)u(x)v(x)\,dx\,. \qquad (35)$$

With the requirements (29) on q and p we see that a is continuous and H_0^1-elliptic. Due to the Lax-Milgram lemma, cf. [4], (34) posesses a unique solution u in H_0^1. We get the n-dimensional linear system

$$A_e\, u_e \;=\; f_e \qquad (36)$$

by an application of a Galerkin method to (34) with respect to the n-dimensional subspace $V_h \subset H_0^1$,

$$V_h \;=\; \mathrm{span}\{\, v_0, \ldots, v_{n-1}\,\}. \qquad (37)$$

The functions v_i are chosen to be piecewise linear on the grid $\mathcal{G}$ with $v_i(x_k) = \delta_{i,k}$. A_e is a positive definite, tridiagonal matrix with entries $(A_e)_{ik} = a(v_i, v_k)$ for $|i-k| \le 1$ and satisfies

$$\underline{q}\, h^{-1}\, \mathcal{D} \;\le\; A_e \;\le\; \overline{q}\, h^{-1}\, \mathcal{D} \;+\; \overline{p}\, h\, \tilde{I} \qquad (38)$$

where $h\,\tilde{I}$ symbolizes the matrix $(v_i, v_k)_0$. f_e has the components $(f_e)_i = (f, v_i)_0$ and $u_h = \sum (u_e)_i v_i$ represents the Finite-Element approximation to the solution u of (34).

Lemma 3.3 *Let* $n \ge N + \left[\frac{n-1}{2}\right]$, *then*

$$(G_N^n)^* G_N^n \;\le\; \frac{1}{2}\, \mathcal{D}\,. \qquad (39)$$

Before we prove the lemma we state some of its obvious implications. The criterion (25) holds for both discretization matrices A_d and A_e.

Corollary 3.4 *Let* $n \geq N + \left[\frac{n-1}{2}\right]$, *then*

$$(G_N^n)^* G_N^n \leq \frac{h^2}{2\,\underline{q}}\, A_d\,, \tag{40}$$

$$(G_N^n)^* G_N^n \leq \frac{h}{2\,\underline{q}}\, A_e\,. \tag{41}$$

Proof: The inequalities (40) and (41) follow directly by the first part of (32) resp. (38) and (39). ♣

Corollary 3.5 *Solving the equations (30) resp. (36) by WTLM with coarse grid correction (23) and damped Jacobi iteration as BIM (i.e.* $L = \vartheta^{-1} D_A$ *in (15)) gives (21) as well as (22) with*

$$C_{A_d} = C_{A_e} = \frac{\overline{q}}{\underline{q}\,\vartheta} \quad \textit{for} \quad 0 < \vartheta \leq \frac{\underline{q}}{2\,\overline{q}}\,.$$

Using the Gauß-Seidel relaxation as BIM results in (22) with

$$C_{A_d} = \frac{\overline{q}}{\underline{q}} \left(1 + \frac{\overline{q}}{2\,\underline{q}}\right)^2,$$

$$C_{A_e} = \left(\frac{\overline{q}}{\underline{q}} + \frac{\overline{p}\,h^2}{3\,\underline{q}}\right) \left(1 + \frac{\overline{q}}{2\,\underline{q}} + \frac{\overline{p}\,h^2}{12\,\underline{q}}\right)^2,$$

respectively.

Proof: We will only calculate the constants C_{A_d} for both BIMs. The constants of the FE case can be obtained similarly. Let us begin with the Jacobi method. First we determine the range of ϑ to assure $\vartheta^{-1} D_{A_d} \geq A_d$ or equivalent $0 < \vartheta \leq \|D_{A_d}^{-1/2}\, A_d\, D_{A_d}^{-1/2}\|^{-1}$. The estimate

$$\begin{aligned}\|D_{A_d}^{-1/2}\, A_d\, D_{A_d}^{-1/2}\| &\leq \|D_{A_d}^{-1}\|\,\|A_d\| \overset{(32)}{\leq} \frac{h^2}{2\,\underline{q}}\,(\overline{q}\,h^{-2} \underbrace{\|\mathcal{D}\|}_{\leq 4} + \overline{p}) \\ &\leq \frac{2\,\overline{q}}{\underline{q}} + \frac{h^2\,\overline{p}}{2\,\underline{q}}\end{aligned}$$

shows that $0 < \vartheta \leq \underline{q}/(2\overline{q})$ is sufficient. So, the conditions of Theorem 3.1 are satisfied and

$$C_{A_d} = \gamma \|L\| \overset{(40)}{\leq} \frac{h^2}{2\,\underline{q}\,\vartheta} \|D_{A_d}\| \leq \frac{h^2}{2\,\underline{q}\,\vartheta} \frac{2\,\overline{q}}{h^2} = \frac{\overline{q}}{\underline{q}\,\vartheta}\,.$$

In the case of the Gauß-Seidel relaxation we apply Lemma 3.2. Therefore we need a bound for $\|I - D_{A_d}^{-1/2}\, F\, D_{A_d}^{-1/2}\|$ where F is the strictly upper triangular part of A_d. Now, $\|F\|^2 = \varrho(F^T F) \leq \|F^T\|_* \|F\|_*$ with the maximal row sum $\|\cdot\|_*$. Hence,

$$\|I - D_{A_d}^{-1/2}\, F\, D_{A_d}^{-1/2}\| \leq 1 + \|D_{A_d}^{-1}\| \, \|F\| \leq 1 + \frac{h^2}{2\,\underline{q}} \frac{q}{h^2} = 1 + \frac{\overline{q}}{2\,\underline{q}}\,,$$

and finally

$$C_{A_d} = \gamma \|D_{A_d}\| \, \|I - D_{A_d}^{-1/2}\, F\, D_{A_d}^{-1/2}\|^2 \overset{(40)}{\leq} \frac{\overline{q}}{\underline{q}} \left(1 + \frac{\overline{q}}{2\,\underline{q}}\right)^2 . \quad \clubsuit$$

Proof of Lemma 3.3: Instead of (39) we verify the equivalent inequality

$$G\,\mathcal{D}^{-1}\,G^* \leq 1/2 \tag{42}$$

in two steps. First, we show

$$\mathcal{D}^{-1} \leq M \quad \text{with} \quad M \in \mathbb{R}^{n \times n}\,, \quad M_{ij} = \min\{i+1, j+1\}\,, \tag{43}$$

and then

$$G\,M\,G^* \leq 1/2\,. \tag{44}$$

We see that (43) and (44) imply (42). Moreover (43) can be reformulated in the form $I \leq \mathcal{D}^{1/2} M \mathcal{D}^{1/2}$ which is true if and only if any eigenvalue of $\mathcal{D}^{1/2} M \mathcal{D}^{1/2}$ is greater or equal to 1. Since $\mathcal{D}^{1/2} M \mathcal{D}^{1/2}$ is regular its spectrum coincide with that one of $M\mathcal{D}$, i.e. $\sigma(\mathcal{D}^{1/2} M \mathcal{D}^{1/2}) = \sigma(M\mathcal{D})$.

By calculating the product $M\mathcal{D}$ which is triangular we derive $\sigma(M\mathcal{D}) = \{1, n+1\}$. For $N-1 \le i, j \le [(n-2)/2]$ we get

$$(G\,M\,G^*)_{ij} \;=\; \sum_{m,l=2-2N}^{1} g_N(l)\, g_N(m)\, \min\{m+2i+1,\, l+2j+1\}\,.$$

For N fixed the sum depends only on $|i-j|$ and $(G\,M\,G^*)_{ij}$ vanishes if $|i-j| \ge N$, i.e. $(G\,M\,G^*)_{ij} = (G\,M\,G^*)_{N-1,N-1+|i-j|} =: r_N(|i-j|)$ with $r_N(z) = 0$, $z \ge N$. If we now take into account the remaining matrix entries $(G\,M\,G^*)_{ij}$, $0 \le i, j \le N-2$, $[(n-2)/2] \le i, j \le n_1 - 1$, and apply the Gershgorin estimate for eigenvalues, cf. Stoer/Bulirsch [22], we have

$$\|G\,M\,G^*\| \;\le\; r_N(0) + 2\sum_{i=1}^{N-1} |r_N(i)| \;=\; B_N\,.$$

Computationally, B_N can be shown to be $1/2$ for all N and so, (44) is established. ♣

Performing WTLM (and later the wavelet multi-level method) in order to solve (30,36) we have to operate with the matrices HA_dH^* and HA_eH^*. Therefore we have a closer look to them.

Lemma 3.6 *If $M \in \mathbb{R}^{n\times n}$, $n \ge N + [(n-1)/2]$, is a b-band matrix, i.e. $M_{ij} = 0$ for $|i-j| > b$, then $H_N^n M (H_N^n)^*$ is a $[(b+2N-1)/2]$-band matrix.*

Proof: The representation

$$(H\,M)_{il} \;=\; \sum_{k=\max\{0,-b+l+2i\}}^{\min\{2N-1,b+l-2i\}} h_N(k)\, M_{k+2i,l}$$

implies $(H\,M)_{il} = 0$ for $l-2i < -b$ or $l-2i > 2N-1+b$. We set $CH^* = (HC^*)^*$ with $C^*_{il} = 0$ for $i-2l < -b$ or $i-2l > 2N-1+b$. As above

$$(HC^*)_{jm} \;=\; \sum_{k=\max\{0,-b-2j+2m\}}^{\min\{2N-1,2N-1+b-2j+2m\}} h_N(k)\, C^*_{k+2j,m}$$

shows $(HC^*)_{jm} = 0$ if $m-j < (1-2N-b)/2$ or $m-j > (2N-1+b)/2$. We finish with $(HMH^*)_{jm} = 0$ for $|m-j| > (2N-1+b)/2$. ♣

Corollary 3.7 *Let A_d and A_e be the 1-band matrices in (30) resp. (36) with $n \ge N + [(n-1)/2]$. Then, $H_N^n A (H_N^n)^*$ is a N-band matrix for $A \in \{A_d, A_e\}$.*

Remark: The Transition $A \to H_N^n A(H_N^n)^*$ increases the band width of A moderately as far as N becomes not too large.

Lemma 3.8 *We adopt the assumptions of the corollary above. Further, we assume* $\overline{p} = 0$*. If* $A \in \{A_d, A_e\}$*, then*

$$cond(H_N^n A(H_N^n)^*) \leq \frac{1}{2} \frac{\overline{q}}{\underline{q}}\, cond(\mathcal{D})\,(1+\varepsilon_n) \tag{45}$$

where $\varepsilon_n = \mathcal{O}(n^{-2})$*.*

Remark: The Transition $A \to H_N^n A(H_N^n)^*$ divides the upper bound of the spectral condition number almost into halves. Also, the relation (45) holds qualitatively for $\overline{p} > 0$.

Proof: We prove $\|\, H\mathcal{D}H^*\| \leq 2$, then (45) follows from (32) resp. (38) and $\|\mathcal{D}\| = 2 - 2\cos(n\pi/(n+1))$. Straightforwardly we obtain, $N-1 \leq i,j \leq [(n-2)/2]$,

$$\begin{aligned}(H\mathcal{D}H^*)_{il} \;&=\; 2\,\delta_{0,i-l} - \sum_{k=0}^{2N-1} h_N(k) \\ &\times \;\; (\,h_N(k+2(i-l)+1) \,+\, h_N(k+2(i-l)-1)\,)\,.\end{aligned} \tag{46}$$

As in the proof of Lemma 3.3 we apply Gershgorin's theorem,

$$\|\, H\mathcal{D}H^*\| \;\leq\; \sum_{k=-N}^{N} |(H\mathcal{D}H^*)_{0k}|\,.$$

The sum over k can be computed to be 2 independently of N. ♣

Remarks:

a) Exept for some special cases it is generally not possible to verify analytically one of the criterions (25) or (26) relative to the matrix A. In view of the computational purpose it may be sufficient to prove one of them only for suitable chosen test vectors. We recommend the following method. According to the interpretation of (25) and (26) we define three test vectors by $(i = 0,1,2)$

$$e^i \;=\; (e_0^{\mu_i}, \ldots, e_{n-1}^{\mu_i}) \quad \text{with} \quad \mu_0 = 0,\; \mu_1 = n/2,\; \mu_2 = n-1\,,$$

$$e_j^{\mu_i} \;=\; \sin\left(\frac{(\mu_i+1)\,\pi\,(j+1)}{n+1}\right)\,, \quad 0 \leq j \leq n-1\,,$$

Table 1: The rates δ_i^H (47) relative to the matrix (48) with filter number $N = 3$ and $\gamma = 1./0.97$

	δ_0^H	δ_1^H	δ_2^H
$n = 1000$	0.9700	0.3553	$5.95 \cdot 10^{-4}$
$n = 1500$	0.9700	0.3557	$5.94 \cdot 10^{-4}$
$n = 2000$	0.9700	0.3560	$5.96 \cdot 10^{-4}$

because they represent the essential frequency domains. By $\gamma > 1/\min\{A_{ii}\}$ we have a reasonable estimate for γ being not too small. Compute now the rates

$$\delta_i^H := \frac{\langle He^i, He^i\rangle}{\gamma\,\langle Ae^i, e^i\rangle}, \qquad \delta_i^G := \frac{\langle Ge^i, Ge^i\rangle}{\gamma\,\langle Ae^i, e^i\rangle} = \frac{1}{\gamma\,\langle Ae^i, e^i\rangle} - \delta_i^H . \tag{47}$$

If $\delta_i^H \leq 1$ for $i = 0, 1$ (and for increasing dimension) then apply the coarse grid correction (24). Conversely, if $\delta_i^G \leq 1$ for $i = 1, 2$ then use the one of (23). The performance of this approach will be demonstrated on the next item.

b) *Outlook on WTLMs applied to integral equations of the first kind*
Generally, the eigenfunctions of an integral operator of the first kind belonging to small eigenvalues are known to be high frequency. This can be expressed in terms of Sobolev norms of high orders. In fact these norms of the eigenfunctions increase as the corresponding eigenvalue decreases, cf. Louis [15]. Consequently, criterion (26) should be satisfied if the matrix originates from the discretization of an integral operator. For example, the matrix

$$A_{ij} = \begin{cases} 0.97 & : \quad i = j \\ \dfrac{\sin(\pi\,(i-j)\,0.97)}{\pi\,(i-j)} & : \quad \text{otherwise} \end{cases} . \tag{48}$$

can be interpreted as a discretization of an integral operator of the first kind and it plays an important rôle in signal analysis, Slepian [21], as well as in computerized tomography, Natterer [17]. So far, there was no success in verifying (26) for that matrix. But the values of δ_i^H in Table 1 indicate that (26) holds, the more so as $\delta_2^G \approx 2 \cdot 10^5$ for $n = 1000$. That is also reflected by the convergence rates of WTLM with coarse grid correction (24), cf.

Table 2: The convergence rates of WTLM with (24) and Gauß-Seidel relaxation relative to the matrix (48) of dimension 200. The convergence rate of the Gauß-Seidel relaxation is 0.97

	$N=1$	$N=2$	$N=3$	$N=4$	$N=5$
$\varrho(\,M^{TL}(1,0)\,)$	0.5994	0.4490	0.0192	0.0189	0.0188

Table 2. Further, the inverse matrices of A_d (30) and A_e (36) fit for this remark due to

$$(H_N^n)^* H_N^n \,\le\, \frac{2\,\overline{q}}{h^2}\,A_d^{-1} \quad \text{and} \quad (H_N^n)^* H_N^n \,\le\, \frac{2\,\overline{q}}{h}\,A_e^{-1}\,.$$

Actually the regularisation property is the right measure to compare iterative schemes applied to the discretization of compact operators rather than the convergence rate, cf. [15]. This short outline should only indicate the principal capacity to treat such linear system with a WTLM due to (26) and Theorem 3.1. Considerable work in this field remains to be done.

4 The wavelet multi-level method

4.1 Some theoretical aspects

The step from the two-level to the multi-level method is straightforward. We start with our original equation (14) and introduce some abbreviations and associations:

$$\begin{aligned}
l &= \min\{\, j \in \mathbb{N}_0 \,:\, n-2N+1 < 2^j \,\} && \text{depth of recursion}\,,\\
k &\in \{\,0,\ldots,l\,\} && \text{level number}\,,\\
n_k &= 2N-1+\left[\frac{n-2N+1}{2^k}\right],\ n = n_0, && \text{dimension on level } k\,,\\
H_k &= H_N^{n_k}\,,\ H = H_0, && \text{restriction form level } k \text{ to level } k+1\,,\\
A_{k+1} &= H_k\,A_k\,H_k^*\,,\ A_0 = A, && \text{matrix on level } k+1\,.
\end{aligned}$$

The matrices A_k are positive definite on all levels. To define the wavelet multi-level method, we first declare the recursive multi-level procedure (MLP). Therefore we denote the unit matrix on $\mathbb{R}^{n_k}$ by I_k and obtain the iteration matrices S_k and the matrices L_k of the BIM, cf. (15), on level k (i.e. relative to A_k) via $S_k = I_k - L_k^{-1}\,A_k$.

MLP(k, l, x, b)
begin
if $k = l$
then

$$x := A_l^{-1} b \qquad \textit{(exact solution on the highest level)} \tag{49}$$

else

$$x := S_k^{\nu_k} x + \sum_{i=0}^{\nu_k - 1} S_k^i L_k^{-1} b \qquad (\ \nu_k \ \textit{steps of BIM on level } k) \tag{50}$$

$$d := H_k (A_k x - b) \qquad \textit{(restriction of the defect to level } k+1\textit{)}$$

$$y := 0$$

$$\textbf{MLP}(k+1, l, y, d) \qquad \textit{(MLP starting on level } k+1 \textit{ with initial guess } y = 0\textit{)} \tag{51}$$

$$x := x - H_k^* y \qquad \textit{(prolongation of the } k+1 \textit{ level approximation to the lower level } k\textit{, coarse grid correction)}$$

end

Remark: MLP describes one V-Cycle with pre-smoothing only. To achieve post-smoothing simply add (50) after the coarse grid correction. Doubling the statement (51) leads to a W-Cycle.

One step of the wavelet multi-level method (WMLM) is performed by

$$\begin{aligned} \tilde{x} &:= x^m \\ &\mathrm{MLP}(0, l, \tilde{x}, b) \\ x^{m+1} &:= \tilde{x} \end{aligned}$$

The only difference to WTLM consists in the statement (51) instead of the exact solution $y := A_{k+1}^{-1} d$.

Formally the above procedure MLP is quite correct but we do not know whether WMLM converges. Even we do not know whether the convergence rate of WTLM relative to intermediate levels is bounded by a constant smaller than 1 independently of the dimension n. Our results of the former section can only be applied to WTLM acting between the levels 0 and 1.

The following two lemmata are supplied to demonstrate that inequalities like (40) and (41) hold on any level. Consequently, similar statements like those of Corollary 3.5 hold for WTLM on all intermediate levels.

Lemma 4.1 *Let $1 \leq N \leq 5$ and $n > 2N$. Let $[\mathcal{D}]_k$ denote the n_k-dimensional matrix corresponding to the differene stencil $[-1 \quad 2 \ -1]$, cf. Section 3.3.1. Then, there exists a positive constant c_N which depends only on N such that*

$$[\mathcal{D}]_k \, < \, c_N \, H_N^{n_{k-1}} \, [\mathcal{D}]_{k-1} \, (H_N^{n_{k-1}})^* \tag{52}$$

holds on all levels $k \in \{1, \ldots, l-1\}$.

Remark: We chose the notation $[\mathcal{D}]_k$ to distinguish clearly from $\mathcal{D}_k = H_{k-1} \, \mathcal{D}_{k-1} \, H_{k-1}^*$.

Proof: Without loss of generality we may assume $k = 1$.
$H\mathcal{D}H^*$ can be split into a sum of a N-band Toeplitz matrix T and a matrix R wich has almost only zero entries except of two $N \times N$ blocks located at the ends of its diagonal,

$$H\mathcal{D}H^* \, = \, T \, + \, R$$

with

$$\begin{aligned} T_{ij} \ &\stackrel{(46)}{=} \ 2\,\delta_{0,i-l} \, - \, \sum_{k=0}^{2N-1} h_N(k) \\ & \qquad \times \quad (\, h_N(k + 2(i-l) + 1) \, + \, h_N(k + 2(i-l) - 1)\,) \end{aligned}$$

and

$$R \, = \, \begin{pmatrix} R^1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & & & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ \vdots & 0 & & 0 & 0 \\ 0 & \cdots & \cdots & 0 & R^2 \end{pmatrix}, \qquad R^i \in \mathbb{R}^{N \times N}.$$

Now, we have $c_N \, H\mathcal{D}H^* - [\mathcal{D}]_1 \, = \, (c_N \, T - [\mathcal{D}]_1) + c_N \, R$. For our further calculations we set $c_N := 10/(\min\{(H\mathcal{D}H^*)_{ii}\} \, \min\{h_N(j)\})$ (In order to hold (52) c_N has to be at least greater than $2/\min\{(H\mathcal{D}H^*)_{ii}\}$). By an application of Lemma 3 of [18], Parter, one has that $\mathcal{T} := c_N \, T - [\mathcal{D}]_1$ is positive definite for all n. Finally, we show

$$0 \, < \, \mathcal{T} \, + \, c_N \, R \, . \tag{53}$$

Because of the simpel structure of R and the positive definitness as well as the Toeplitz form of $\mathcal{T}$ (53) follows by

$$0 < \mathcal{T}^i + c_N R^i \qquad \text{for} \quad i = 1,2 \tag{54}$$

where $\mathcal{T}^i$ denotes these submatrices of $\mathcal{T}$ corresponding to the submatrices R^i of R, i.e. $R^i = (R)^i$. However, (54) is equivalent to

$$0 < (c_N \, H \mathcal{D} H^* - [\mathcal{D}]_1)^i \qquad \text{for} \quad i = 1,2\,.$$

These are relations for $N \times N$ matrices which can be checked numerically either by a Cholesky decomposition or by computing the eigenvalues. We have done both for N up to 5. ♣.

Remark: The statement of the above lemma is limited to filter numbers smaller or equal to 5. This is no principal bound. To extend the lemma to higher N's one has to check the positive definitness of $c_N T - [\mathcal{D}]_1$ as well as the relations (54) with an appropriate choice for c_N.

Lemma 4.2 *We adopt the assumptions and notations of the former lemma. Further, let A_d and A_e be the matrices of (30) resp. (36) corresponding to the discretization step size $h = 1/(n+1)$. Then,*

$$(G_N^{n_k})^* G_N^{n_k} \;\leq\; \frac{h^2}{2\,\underline{q}}\, c_N^k \, (A_d)_k\,, \tag{55}$$

$$(G_N^{n_k})^* G_N^{n_k} \;\leq\; \frac{h}{2\,\underline{q}}\, c_N^k \, (A_e)_k\,, \tag{56}$$

for $k = 0, \ldots, l-1$.

Proof: We apply (52) k-times:

$$\begin{aligned}
G_k^* G_k \;&\overset{(39)}{\leq}\; \frac{1}{2}\,[\mathcal{D}]_k \\
&\leq\; \frac{c_N}{2}\, H_{k-1}\,[D]_{k-1}\, H_{k-1}^* \\
&\leq\; \frac{c_N^k}{2}\, H_{k-1}\cdots H_0\,\mathcal{D}\, H_0^* \cdots H_{k-1}^* \\
&\overset{(32)}{\leq}\; \frac{c_N^k\, h^2}{2\,\underline{q}}\, \underbrace{H_{k-1}\cdots H_0\, A_d\, H_0^* \cdots H_{k-1}^*}_{=\,(A_d)_k}\,.
\end{aligned}$$

By (38) the last estimate may be replaced to yield (56). ♣

Hence, criterion (25) is satisfied on any intermediate level, i.e. the use of the prolongation H^* and the restriction H on any level is now justified. Unfortunately, by no means a bounded convergence rate of a two-level method relative to all intermediate levels implies a bounded rate for the corresponding multi-level method. The convergence rate of the multi-level method may depend on the number of levels. An example is given in [20] which is exactly our WMLM for $N = 1$ applied to the matrix $\mathcal{D}$. However the authors have also pointed out that prolongations which interpolate smooth functions quite well lead to the desired multi-level convergence behaviour. In view of the interpolation result for the finite dimensional wavelet operators, Lemma 2.5, it is not unreasonable to predict a convergence behaviour of WMLM, $2 \leq N \leq 5$, similar to usual multi-grid convergence for solving the linear systems (30) and (36). Note that Lemma 2.5 does not hold for $N = 1$.

4.2 Numerical experiments

We shall provide numerical approximations to the *rate of convergence*, i.e. the spectral radius of the iteration matrix, for both, the WTLM as well as the usual multi-grid method (MGM) relative to the matrix A_d (31) for $p \equiv 0$ and several functions q (27). The convergence rates relative to the FE matrix A_e (36) are relinquished because approximating the integrals which define the entries of A_e by a suitable chosen quadrature rule produces (up to a factor $1/h$) the matrix A_d as far as $p \equiv 0$ and q is sufficiently smooth. Quadrature rules of higher order would also yield comparable rates to those for A_d.
At all following experiments the approximations r to the convergence rates are computed by

$$r = \left(\frac{\|x^{m^*} - x^{m^*-1}\|}{\|x^1 - x^0\|} \right)^{1/(m^* - 1)} \tag{57}$$

where $\{x^m\}_{m \in \mathbb{N}_0}$ denotes the sequence generated by either WMLM or MGM with initial guess $x^0 = (\,(-1)^i \mid 0 \leq i \leq n_0 - 1\,)^T / \sqrt{n_0}$. Further, we always fixed m^* to be 20 and the right hand side f_d of (30) to be 0. For both multi-level iterations we took the Gauß-Seidel relaxation as BIM. In connection with MGM we used the standard piecewise linear interpolation for the prolongation and its 'weighted' adjoint as restriction, cf. [11] (2.3.7b)

Table 3: The rates r (57) for WMLM ($N = 4$) and MGM relative to the matrix A_d (31) with $q_1 \equiv 1$ and $p \equiv 0$

n_0	511		1023	
ν	1	2	1	2
WMLM	0.3330	0.1161	0.3332	0.1156
MGM	0.3380	0.1267	0.3383	0.1245

resp. (2.3.9b). The finest discretization step size h, cf. (31), was adapted for MGM to work perfectly well, i.e. $h = 2^{-j} = 1/(n_0+1)$, $j \in \mathbb{N}$ (without loss of generality we assumed $\alpha = 0$ and $\beta = 1$ in (27)). In order not to confuse the reader by too many numbers we considered WMLM only with the wavelet filter $N = 4$. The presented results differ not very much for the filters $N = 3$ or $N = 5$, but for $N = 2$ the convergence rates would be a little bit worse.
Instead of terminating the recursion of MLP on level l (49), we solved exactly the equation on the first level $k^* < l$ with $n_{k^*} \leq 25$. The reason is that the dimensions n_k of all levels larger than k^* decrease too slowly with increasing k. In the same fashion we treated MGM. The numbers ν which appear in the tables indicate the numbers of the (pre-)smoothing steps on each level, in other words $\nu_k \equiv \nu$ in (50).

The first example deals with the constant coefficient $q_1 \equiv 1$. The values of r can be seen in Table 3 and reflect the typical multi-level convergence behaviour: increasing the dimension does not effect the convergence rate, while doubling ν produces exactly that shrinking of r predicted by the theory [11]. Looking only at the convergence rates there is no difference between MGM and WMLM. Thus far, we have not brought the issue of effort into the discussion of the comparision. In the preparation phase of WMLM we have to compute the lower dimensional matrices $A_1, \ldots, A_{k^*}$ and to store them. The computational work for that is almost equivalent to N iteration steps of MGM. Further, the prolongations and restrictions of WMLM ($N \geq 2$) are a little bit more expensive than these of MGM. Hence, MGM is more efficient all things considered. On the other hand MGM lacks in efficiency if the discretization step size has not the special form $h = 2^{-j}$, $j \in \mathbb{N}$, or the differential equation (27) is dicretized with respect to an irregular grid. Moreover, its prolongations and restrictions have to be carefully adapted to the new situation. A change of h or an irregular grid affect neither the code

Table 4: The rates r (57) for WMLM ($N=4$) and MGM relative to the matrix A_d (31) with $q_2(x) = 1.01 + \sin(4\pi x)$ and $p \equiv 0$

n_0	511		1023	
ν	1	2	1	2
WMLM	0.5166	0.4426	0.5044	0.4118
MGM	0.4697	0.4431	0.4248	0.4424

nor the performance of WMLM.

Next we demonstrate both methods by the more extreme coefficient function $q_2(x) = 1.01 + \sin(4\pi x)$ which is still smooth but oscillating and the rate $\overline{q}/\underline{q}$ of its extrem values is 201. Due to Corollary 3.5 and Theorem 3.1 we know how the convergence rate of the WTLM gets worse by an increase of $\overline{q}/\underline{q}$. In the quite same way WMLM as well as MGM behave, cf. Table 4. Last we consider two, *only slightly different*, discontinuous coefficients

$$q_3(x) = \begin{cases} 0.1 & : \quad x \le 0.5 \\ 10 & : \quad \text{otherwise} \end{cases} \quad \text{and} \quad q_4(x) = \begin{cases} 0.1 & : \quad x \le 0.51 \\ 10 & : \quad \text{otherwise} \end{cases} . \tag{58}$$

Table 5 is arranged to bring the convergence rates of the particular methods for q_3 face to face with the rates for q_4. At a first glance it is surprising that MGM works very well for q_3 but breaks completely down for q_4. A few words are in order: The point of discontinuity of q_3 is just a grid point of all grids which are involved in MGM. In this case it is well known that no problems arise during the multi-grid solution process. However, if this favourable fact is not met one has to introduce a problem(matrix)-dependent prolongation and restriction, cf. [11] Chapter 10.3, which save the multi-grid process, but make it more expensive. We did not change the prolongations and restrictions in MGM and that explains the high values of r in Table 5 with respect to q_4. Certainly, the rates of WMLM for q_3 are not impressive compared to those of MGM, but the convergence behaviour of WMLM is much more stable under a perturbation of the point of discontinuity. We chose this example to show the relative robustness of WMLM for the so called *interface problems*. The differences of MGM and WMLM base on the manner how the prolongation resp. restriction operators are designed. In the multi-grid theory the construction of these operators is motivated by an interpolation viewpoint and so they have to be adapted if the 'smoothness' of the problem changes. Whereas their choice in WMLM underlies a kind of

Table 5: The rates r (57) for WMLM ($N = 4$) and MGM relative to the matrix A_d (31) with q_3 as well as q_4, see (58), $p \equiv 0$

	n_0	511		1023	
	ν	1	2	1	2
WMLM	q_3	0.7654	0.6320	0.7698	0.6629
	q_4	0.7372	0.6971	0.7227	0.6949
MGM	q_3	0.3395	0.1298	0.3391	0.1275
	q_4	1.7626	1.5295	2.8486	1.6091

spectral equivalence on all levels, cf. (25) and (26), which holds for a larger class of problems: (55) and (56) are still valid for discontinuous coefficients in the differential equation (27).

References

[1] Battle G., Federbush P.: *Ondelettes and Phase Cell Cluster Expansions. A Vindication*, Commun. Math. Phys., 109, 1987, p. 417-419

[2] Beylkin G., Coifman R., Rokhlin V.: *Fast Wavelet Transforms and Numerical Algorithms I*, Commun. Pure and Appl. Math., Vol. XLIV, 1991, p. 141-183

[3] Briggs W. L., Henson V. E.: *Wavelets and Multigrid*, SIAM J. Sci. Comput., 14(2), 1993, p. 506-510

[4] Ciarlet G. C.: *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1980

[5] Daubechies I.: *Othonormal Bases of Compactly Supported Wavelets*, Commun. Pure and Appl. Math., Vol. XLI, No.7, 1988, p. 909-996

[6] Daubechies I.: *The Wavelet Transform, Time-Frequency Localisation and Signal Analysis*, IEEE Trans. Inform. Theory, vol. 36, No. 5, 1990, p. 961-1005

[7] Daubechies I.: *Ten Lectures on Wavelets*, SIAM Publ., Philadelphia, 1992

[8] DeVore R., Jawerth B., Lucier B.: *Image compression through wavelet transform coding*, IEEE Trans Inform. Theory, vol. 38, 1992, p. 719-746

[9] Glowinski R., Rieder A., Wells R. O., Zhou X.: *Wavelet multigrid for Dirichlet boundary value problems in domains with general shape*, Technical Report of the Computational Mathematics Laboratory, Rice University, Houston, to appear in 1993

[10] Grossmann A., Holschneider M., Kronland-Martinet, Morlet J.: *Detection of abrupt signal changes with the help of the wavelet transform*, Reprint, Centre de Physique théorique, C.N.R.S. Luminy Case 907, 13288 Marseille Cedex 9, France

[11] Hackbusch W.: *Multi-Grid Methods and Applications*, Springer Series in Computational Mathematics 4, Springer-Verlag Berlin, 1985

[12] Hackbusch W.: *The Frequency Decomposition Multi-Grid Method, Part I: Application to Anisotropic Equations*, Numer. Math. 56, p. 229-245, 1989

[13] Heil C. E., Walnut D. F.: *Continuous and Discrete Wavelet Transforms*, SIAM Review, 31(4), 1989, p. 628-666

[14] Holschneider M.: *On the Wavelet Transformation of Fractal Objects*, J. Stat. Phys. 50, 1988, p. 963-993

[15] Louis, A. K.: *Inverse und schlechtgestellte Probleme*, Teubner Studienbücher Mathematik, Teubner Stuttgart, 1989

[16] Meyer Y.: *Ondelettes*, Hermann Paris, 1990

[17] Natterer F.: *The Mathematics of Computerized Tomography*, Wiley-Teubner, New York, 1987

[18] Parter S. V.: *Extreme Eigenvalues of Toeplitz Forms and Applications to Elliptic Difference Equations*, Trans. Amer. Math. Soc. 99, 1961, p. 153-193

[19] Rieder A.: *Wavelet Transform on Sobolev Spaces and its Approximation Properties*, Numer. Math. 58, 1991, p. 875-894

[20] Ruge J. W., Stüben K.: *Algebraic Multigrid*, in: *Multigrid Methods*, S. F. McCormick ed., SIAM, Philadelphia, 1987

[21] Slepian D.: *Prolate Spheriodal Wave Functions, Fourier Analysis and Uncertainty, V: The Discrete Case*, Bell Syst. Techn. J., 57, 1978, p. 1371-1430

[22] Stoer J., Bulirsch R.: *Introduction to Numerical Analysis*, Springer, New York, 1980

[23] Strang G.: *Wavelets and Dilation Equations: A Brief Introduction*, Siam Review, 31(4), 1989, p. 614-627

[24] Stüben K.: *Algebraic multi-grid (AMG): experiences and comparisions*, Appl. Math. Comput., 13, 1983, p. 419-451

[25] Varga R. S.: *Matrix Iterative Analysis*, Prentice-Hall Inc., 1962

[26] Wells R. O., Zhou X.: *Wavelet Solutions for the Dirichlet Problem*, Technical Report of the Computational Mathematics Laboratory, No. CML TR92-02, Rice University, Houston, 1992

[27] Xu J.-C., Shann W.-C.: *Galerkin–Wavelet Methods for Two–Point Boundary Value Problems*, Numer. Math. 63, 123-144(1992)