Wavelet-discretizations of parabolic integro-differential equations

Author(s):
Petersdorff, T. von; Schwab, Christoph

Publication Date:
2001

Permanent Link:
https://doi.org/10.3929/ethz-a-004289317

Wavelet-discretizations of parabolic integro-differential equations*

T. von Petersdorff and C. Schwab

Research Report No. 2001-07
August 2001

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

*Research supported in part by the TMR-network "Wavelets and Multiscale methods in numerical analysis" under BBW project number 97.404

Wavelet-discretizations of parabolic integro-differential equations*

T. von Petersdorff and C. Schwab

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2001-07 August 2001

Abstract

We consider parabolic problems $\dot{u}+Au=f$ in $(0,T)\times\Omega$, $T<\infty$, where $\Omega\subset\mathbb{R}^d$ is a bounded domain and A is a strongly elliptic, classical pseudo-differential operator of order $\rho\in[0,2]$ in $\tilde{H}^{\rho/2}(\Omega)$. We use a θ-scheme for time discretization and a Galerkin method with N degrees of freedom for space discretization. The full Galerkin matrix for A can be replaced with a sparse matrix using a wavelet basis, and the linear systems for each time step are solved approximatively with GMRES. We prove that the total cost of the algorithm for M time steps is bounded by $O(MN(\log N)^\beta)$ operations and $O(N(\log N)^\beta)$ memory. We show that the algorithm gives optimal convergence rates (up to logarithmic terms) for the computed solution with respect to L^2 in time and the energy norm in space.

*Research supported in part by the TMR-network "Wavelets and Multiscale methods in numerical analysis" under BBW project number 97.404

1 Introduction

Fast algorithms such as wavelets, multipole or clustering methods for the numerical solution of elliptic integrodifferential equations

$$A[u](x) = \int_\Omega k(x, x-y)u(y)dy = f \qquad x \in \Omega \tag{1.1}$$

with kernel function $k(x,z)$ have been introduced and analyzed in recent years (see, e.g., [3, 4, 7]). In the present paper we investigate the numerical solution of a class of parabolic integrodifferential equations

$$u_t = A[u](x) + f \quad \text{ in } (0,T) \times \Omega, \tag{1.2}$$

with suitable initial and boundary conditions. Such equations arise as Kolmogorov forward equations for Lévy processes X_t with infinitesimal generators $A[u]$. Brownian motion B_t with diffusion $\sigma(x)$ and drift $r(x)$ is a particular Lévy process. The infinitesimal generator of B_t in dimension $d=1$ is the second order elliptic differential operator

$$A_B[u](x) = -\frac{d}{dx}\left(\sigma(x)\frac{du}{dx}(x)\right) + r(x)\frac{du}{dx} \tag{1.3}$$

and the Kolmogorov forward equation is the diffusion equation with drift. The decomposition theorem of P. Lévy states that the infinitesimal generator A of any Lévy-process X_t is the sum of a differential operator A_B as in (1.3) which accounts for the diffusion part of X_t and could possibly vanish and of a nonlocal operator A_L of the form (1.1) which corresponds to the pure jump part of the process (see, e.g., [2, 12]). The order ρ of the infinitesimal generator A of a Lévy-process always satisfies

$$0 \le \rho \le 2. \tag{1.4}$$

We emphasize that due to (1.4) the kernels $k(x,z)$ are not integrable near $z=0$ and that the integral in (1.1) has to be understood as finite part or principal value, i.e. in the sense of distributions [14]. Interpretations of the integral operators A in the distribution sense can naturally be accounted for in Galerkin discretizations.

While the initial-boundary value problems (1.2) with (1.3) and constant σ, r can be solved analytically for certain initial conditions, numerical solutions are required for nonconstant coefficients, general Lévy processes and free boundary problems arising with optimal stopping of X_t. In a numerical solution, $u(x,t)$ is approximated by Finite Differences or Finite Elements in x with N degrees of freedom, reducing (1.2) to a system of N ordinary differential equations for the approximation u_N which must be integrated in t by a time-stepping scheme. We consider the θ-scheme for time discretization which includes as special cases forward Euler ($\theta=0$), backward Euler ($\theta=1$), and Crank-Nicolson ($\theta=\frac{1}{2}$). In general this leads to implicit methods where a linear system has to be solved for each time step. For the differential operator (1.3) in dimension $d=1$, the matrices to be inverted in each time-step are banded and can be factored in $O(N)$ operations. If the operator A is nonlocal, however, standard Galerkin discretizations of u with N degrees of freedom entail dense stiffness matrices and hence at least $O(N^2)$ complexity per time step for the numerical solution of (1.2). We reduce this complexity by a wavelet-based matrix compression. The basic idea behind this compression is to represent the Galerkin approximation u_N of (1.2) in a wavelet basis. Wavelet matrix compression exploits that the generators A are often classical pseudodifferential operators which implies special properties of their Schwartz kernel function $k(x,z)$ such as

$$\operatorname{sing\,supp}(k(x,z)) \subset \{z=0\}$$

and even analyticity of $k(x,z)$ off the origin $z=0$. Wavelet matrix compression requires only finite differentiability of $k(x,z)$ for $z \neq 0$ and allows the generation of an approximate stiffness matrix of the nonlocal operator A in (1.1) in $O(N(\log N)^a)$ memory and operations where $a \geq 0$ is a small integer (see e.g. [3, 4, 10, 9, 11, 13] and the references there).

The analysis of the impact of this truncation error on stability and consistency of the θ time-stepping scheme for the nonlocal parabolic initial boundary value problems (1.1), (1.2) is the purpose of the present paper. A large body of literature on time-stepping for parabolic problems with Galerkin discretization is available, see [16]. However, our setting with integral operators and matrix compression causes consistency errors which don't fit readily into existing error estimates.

As it is well-known, the stability of explicit time-stepping schemes for Galerkin discretizations for parabolic problems (1.2), (1.4) requires a CFL condition which, as we will show, depends on the order ρ of the operator A and which takes the form

$$\Delta t \leq C(\Delta x)^{\rho}, \quad \rho \in [0,2]. \tag{1.5}$$

For $\rho = 2$, e.g. the heat equation, we recover the classical CFL condition which forces small time-steps Δt in explicit schemes when the meshwidth $h = \Delta x$ of the space discretization is reduced. If, however, the order of A is $\rho \leq 1$, condition (1.5) is of the type usually encountered in time-stepping for first order hyperbolic equations and explicit time-stepping schemes appear competitive.

Next, we present classes of spline wavelets and a matrix compression strategy which leads to sparse approximations for the stiffness matrix of A with $O(N \log N)$ (rather than $O(N^2)$ for standard Galerkin schemes) nonvanishing entries. We prove that this compression preserves the asymptotic convergence rates of the full Galerkin scheme.

In the θ-scheme, a linear system of equations at each implicit time-step must be solved. Since the compressed matrices are not banded and possibly nonsymmetric (due to the presence of a drift term or if $k(x,z)$ is asymmetric for $z \to \pm\infty$), we propose inexact equation solution by GMRES iteration. Using wavelets, we precondition the compressed matrix in dependence on the discretization parameters and the order ρ of A. We relate the GMRES stopping criterion to the discretization error of the scheme and prove that the resulting method converges still of optimal order in space and time while its complexity is essentially $O(N)$ memory and operations per (explicit or implicit) time-step. This is comparable to the complexity for the heat equation using backward Euler in time and banded matrices in space.

We emphasize that our analysis is applicable to general kernels $k(x,z)$, translation invariant or not, of any order $\rho \geq 0$. Therefore, the θ-scheme with wavelet compression allows the numerical solution of the Kolmogorov equations (1.2) for a large class of Lévy processes with complexity comparable to standard finite differences for the heat equation in one dimension.

The outline of the paper is as follows: In Section 2 we present the class of parabolic problems and the class of spatial integro-differential operators A admissible in our analysis. In Section 3, we discuss the fully discrete θ-scheme. We describe wavelet Galerkin discretization of A and give several examples of wavelets. Section 4 is devoted to the stability analysis of the θ-scheme with compression in the "explicit" case $0 \leq \theta < \frac{1}{2}$ as well as in the implicit case $\theta > \frac{1}{2}$. In Section 5, we prove our convergence estimates with particular attention to the error due to wavelet compression of the stiffness matrix $\mathbf{A}$ of A. Section 6 is devoted to the proof of the complexity estimates, the matrix preconditioning in the implicit time-stepping schemes and the analysis of the error in the presence of incomplete GMRES iterations. Throughout, C will denote a generic positive constant independent of the discretization parameters taking different values in different places. If the value of C is relevant, we write also C_i.

2 Problem Formulation

In the time interval $J = (0, T)$ with $T > 0$, we consider parabolic evolution problems of the form

$$u'(t) + Au(t) = g(t),\ t \in J \tag{2.1}$$

$$u(0) = u_0 \tag{2.2}$$

where A is a possibly nonlocal operator of order $\rho > 0$.

For a variational formulation of this problem we introduce Sobolev spaces. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary $\Gamma = \partial\Omega$. We denote by $H = L^2(\Omega)$ the usual square integrable functions with inner product $(.,.)$ and by $H^s(\Omega)$, $s \geq 0$, the corresponding Sobolev spaces (see, e.g., [1]). Further, for $s \geq 0$, we define the space

$$\tilde{H}^s(\Omega) = \left\{ u|_\Omega \mid u \in H^s(\mathbb{R}^d),\, u|_{\mathbb{R}^d \setminus \Omega} = 0 \right\}. \tag{2.3}$$

If $s + 1/2 \notin \mathbb{N}$, then $\tilde{H}^s(\Omega)$ coincides with $H_0^s(\Omega)$, the closure of $C_0^\infty(\Omega)$ with respect to the norm in $H^s(\Omega)$. We identify $L^2(\Omega)$ with its dual and denote by

$$V = \widetilde{H}^{\rho/2}(\Omega). \tag{2.4}$$

Then $V \overset{d}{\hookrightarrow} L^2(\Omega)$ with dense injection and V^*, the dual of V, satisfies

$$V \overset{d}{\hookrightarrow} L^2(\Omega) \overset{d}{\hookrightarrow} V^*. \tag{2.5}$$

We assume that $A \in \mathcal{L}(V, V^*)$. By $(\cdot,\cdot)_{V^*\times V}$ we denote the extension of $(.,.)$ as duality pairing in $V^* \times V$ and by $\|\cdot\|$, $\|\cdot\|_V$, $\|\cdot\|_{V^*}$ the norms in $L^2(\Omega), V, V^*$, resp. We associate with A the bilinear form $a(\cdot,\cdot)$: $V \times V \to \mathbb{C}$ via

$$a(u, v) := (Au, v)_{V^*\times V},\ u, v \in V,\ Au \in V^*\ , \tag{2.6}$$

Then the form $a(\cdot,\cdot)$ is continuous

$$\forall u, v \in V: \quad |a(u, v)| \leq \alpha \|u\|_V \|v\|_V \tag{2.7}$$

and we assume that it is coercive in the sense that

$$\forall u \in V: \quad a(u, u) \geq \beta \|u\|_V^2\ , \tag{2.8}$$

for some $0 < \beta \leq \alpha < \infty$. Then $A \in \mathcal{L}(V, V^*)$ is an isomorphism and $\|A\|_{\mathcal{L}(V,V^*)} \leq \alpha$, $\|A^{-1}\|_{\mathcal{L}(V^*,V)} \leq \frac{1}{\beta}$. The time derivative $u'(t)$ in (2.1) is understood in the weak sense, i.e. for $u \in L^2(J, V)$ we have $u' \in L^2(J, V^*)$ defined by

$$\int_J (u'(t), v)_{V^*\times V}\, \varphi(t) dt = -\int_J (u(t), v)\, \varphi'(t) dt \tag{2.9}$$

for every $v \in V$, $\varphi \in C_0^\infty(J)$. The weak form of (2.1), (2.2) reads:

given

$$u_0 \in H,\ g \in L^2(J;H)\,, \tag{2.10}$$

find $u \in L^2(J,V) \cap H^1(J,V^*)$ such that $u(0) = u_0$ and, for every $v \in V$, $\varphi \in C_0^\infty(J)$,

$$-\int_J (u(t),v)\,\varphi'(t)dt + \int_J a(u,v)\,\varphi(t)dt = \int_J (g(t),v)_{V^*\times V}\,\varphi(t)dt\,. \tag{2.11}$$

Note that the initial condition is well defined since

$$L^2(J,V) \cap H^1(J,V^*) \subset C^0([0,T];\ H)\,. \tag{2.12}$$

Under the assumption (2.10), problem (2.11) has a unique weak solution $u(t)$ and there holds the a-priori estimate [8]

$$\|u\|_{C(\overline{J};H)} + \|u\|_{L^2(J;V)} + \|u'\|_{L^2(J;V^*)} \le C(\|g\|_{L^2(J;H)} + \|u_0\|_H)\,. \tag{2.13}$$

Remark 2.1.

i) We do not assume A to be self-adjoint. The form $a(\cdot,\cdot)$ need not be symmetric.

ii) Properties (2.7) and (2.8) allow to define on V an equivalent norm by

$$\|u\|_a := (a(u,u))^{1/2} \sim \|u\|_V \tag{2.14}$$

to which we shall refer below as "energy-norm".

iii) Testing (2.1) with $u(t)$ in the $(\cdot,\cdot)$ inner product, we find with (2.6) that for almost every $t \in (0,T)$

$$(u,u') + a(u,u) = (u,g)\,,$$

and integrating from $t=0$ to $t=T$, we find

$$\begin{aligned}
&\tfrac{1}{2}\|u(T)\|^2 - \tfrac{1}{2}\|u(0)\|^2 + \int_0^T a(u(t),\,u(t))dt = \int_0^T (u,g)dt \\
&\le \int_0^T \|u(t)\|_a \sup_{v\in V}\frac{(v,g)}{\|v\|_a}\,dt \le \frac{1}{2}\int_0^T \|u(t)\|_a^2\,dt + \frac{1}{2}\int_0^T \|g(t)\|_{V^*}^2\,dt
\end{aligned}$$

which implies the a-priori estimate

$$\|u(T)\|^2 + \int_0^T \|u(t)\|_a^2\,dt \le \|u(0)\|^2 + \int_0^T \|g(t)\|_{V^*}^2\,dt \tag{2.15}$$

where we have set, for any $g \in V^*$,

$$\|g\|_{V^*} = \sup_{v\in V}\frac{(g,v)}{\|v\|_a}\,.$$

Some examples follow.

Example 2.2. (Diffusion problem) Here $\rho = 2$ and

$$A = -\nabla \cdot \mathbf{D}(x)\nabla, \;\; V = H_0^1(\Omega) \overset{d}{\hookrightarrow} L^2(\Omega) = H, \;\; a(u,v) = \int_\Omega \nabla v \cdot \mathbf{D}(x)\nabla u \, dx$$

where $\mathbf{D} \in L^\infty(\Omega)^{d\times d}$ satisfies for some $\gamma > 0$

$$\xi^T \mathbf{D}(x)\xi \geq \gamma\,|\xi|^2 \quad \forall \underline{\xi} \in \mathbb{R}^n, \text{ a.e. } x \in \Omega\,.$$

Then (2.1), (2.2) is the Dirichlet problem for the heat equation in $\Omega \times (0,T)$.

In this example, the operators A are differential operators and in particular local. The nonlocal operators A of interest to us are classical pseudodifferential operators.

Example 2.3. For $0 \leq \rho \in \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ open, bounded and Lipschitz, we consider classical pseudo-differential operators of order $\rho \in [0,2]$ in Ω, i.e. $A \in \Psi^\rho(\Omega)$ which acts from $V \to V^*$ where $V = \widetilde{H}^{\frac{\rho}{2}}(\Omega)$. By the Schwartz kernel theorem (see, e.g. [14]), $A \in \Psi^\rho(\Omega)$ has a representation in terms of a distributional kernel

$$k(x, x-y) \in \mathcal{D}'(\Omega \times \Omega) \tag{2.16}$$

with associated bilinear form

$$a(u,v) = (Au, v)_{V^*\times V} = \big\langle k(x,x-y),\; v(x) \otimes u(y)\big\rangle\,. \tag{2.17}$$

Moreover, the kernel $k(x, x-y) \in C^\infty(\Omega \times \Omega \backslash \{x = y\})$ satisfies the so-called **Calderón-Zygmund estimates**: $\forall \alpha, \beta \in \mathbb{N}_0^n$, $\forall (x,y) \in \Omega \times \Omega\backslash\{x=y\}$:

$$|\partial_x^\alpha\, \partial_y^\beta\, k(x,x-y)| \leq C(\alpha,\beta)|x-y|^{-(d+\rho+|\alpha|+|\beta|)}\,. \tag{2.18}$$

A particular example for a nonlocal operator of order $\rho = 1$ is given by $\Omega = (-1/2, 1/2) \subset \mathbb{R}$ and, for $u \in V = \widetilde{H}^{1/2}(\Omega)$:

$$(Au)(x) = -\text{p.f.} \int_\Omega \frac{u(y)}{|x-y|^2}\, dy \tag{2.19}$$

where the integral is to be understood in the finite-part sense (see, e.g. [14]). For the bilinear form $a(u,v)$ corresponding to the hypersingular operator W in (2.19) integration by parts yields the representation

$$\forall u, v \in \widetilde{H}^{1/2}(\Omega):\; a(u,v) = -\int_\Omega v'(x) \int_\Omega \log(x-y)\, u'(y)\, dy\, dx \tag{2.20}$$

and one can show that there are $\beta, \gamma > 0$ with

$$\forall u \in \widetilde{H}^{1/2}(\Omega):\; a(u,u) \geq \beta\,\|u\|^2_{\widetilde{H}^{1/2}(\Omega)}\,. \tag{2.21}$$

Remark 2.4. In the setting (2.17), we often do not have the coercivity (2.8), but rather a (weaker) **Gårding inequality**: there is $\gamma \geq 0$ such that

$$\forall u \in V:\; a(u,u) + \gamma\,\|u\|^2 \geq \beta\,\|u\|_V^2\,. \tag{2.22}$$

This case can be reduced to (2.8) by the substitution $w = \exp(-\gamma t)u$, since then (2.1) implies that w solves the problem

$$w' + (A + \gamma I)\, w = \exp(-\gamma t)\, g \;\text{ in }\; (0,T)$$

and the operator $A + \gamma I$ is, by (2.19), once again coercive.

3 Discretization

We discretize (2.1) in time using the so-called θ-scheme and in space by a finite element method. We describe wavelet finite element bases and the compression of the stiffness matrix.

3.1 Space Discretization

To discretize the parabolic problem (2.11) in space, we use an elliptic projection onto a family $\{V_h\}_h \subset V$ of finite dimensional subspaces of V, based on piecewise polynomials of degree $p \geq 0$ on a quasiuniform family of triangulations $\{\mathcal{T}_h\}_h$ of Ω.

The semidiscrete problem reads: given $u_0 \in H$, $g \in L^2(J,H)$, first choose an approximation $u_{0,h} \in V_h$ for the initial data u_0. Then find $u_h \in H^1(J, V_h)$ such that

$$u_h(0) = u_{0,h} \tag{3.1}$$

and

$$\frac{d}{dt}(u_h, v_h) + a(u_h, v_h) = (g(t), v_h) \quad \forall v_h \in V_h. \tag{3.2}$$

The approximation of the intial data could be chosen as $u_{0,h} = P_h u_0$ with some projector $P_h\colon L^2 \to V_h$, or as an interpolation of u_0.

The semidiscrete problem (3.1), (3.2) is an initial value problem for $N = \dim V_h$ ordinary differential equations

$$\mathbf{K}\frac{d}{dt}\underline{u} + \mathbf{A}\underline{u} = \underline{g}(t), \quad \underline{u}(0) = \underline{u}_0$$

where $\underline{u}(t)$ denotes coeffcient vector of $u_h(t)$ with respect to some basis of V_h. Likewise $\underline{u}_0$ denotes the coefficient vector of $\underline{u}_0$ and $\mathbf{K}, \mathbf{A}$ denote the mass- and stiffness matrix, respectively, with respect to the basis of V_h.

In the ensuing error analysis, we need to consider functions in V which have additional regularity and introduce for this purpose the spaces $\mathcal{H}^s(\Omega)$ which are defined as

$$\mathcal{H}^s(\Omega) = \begin{cases} V = \tilde{H}^{\rho/2}(\Omega) & \text{for } s = \rho/2, \\ V \cap H^s(\Omega) & \text{for } s > \rho/2. \end{cases}$$

We assume the **approximation property**: for all $u \in \mathcal{H}^t$ with $t \geq \rho/2$ there exists a $u_h \in V_h$ such that for $0 \leq s \leq \frac{\rho}{2}$ and $\rho/2 \leq t \leq p+1$

$$\|u - u_h\|_{\tilde{H}^s(\Omega)} \leq c h^{t-s} \|u\|_{\mathcal{H}^t(\Omega)} \tag{3.3}$$

In section 3.4 below we will use a wavelet basis to define a projection $P_h\colon V \to V_h$ such that $u_h = P_h u$ satisfies (3.3).

We shall also need the **inverse property**: there is $c > 0$ independent of h such that

$$\forall u_h \in V_h \quad \|u_h\|_{\tilde{H}^s(\Omega)} \leq c h^{-s} \|u_h\|_{L^2(\Omega)}, \quad 0 \leq s \leq \rho/2. \tag{3.4}$$

3.2 Time discretization using the θ-scheme

For $T < \infty$ and $M \in \mathbb{N}$, define the time step

$$k = T/M$$

and $t^m = mk$, $m = 0, \ldots, M$. The fully discrete θ-scheme reads as follows: given $u_0 \in H$, find $u_h^m \in V_h$ satisfying

$$u_h^0 = u_{0.h}, \tag{3.5}$$

and, for $m = 0, 1, \ldots, M-1$, find $u_h^{m+1} \in V$ such that for all $v_h \in V_h$

$$\Big(\frac{u_h^{m+1} - u_h^m}{k}, v_h\Big) + a(u_h^{m+\theta}, v_h) = \big(g(t^{m+\theta}), v_h\big) \tag{3.6}$$

holds. Here $u_h^{m+\theta} := \theta u_h^{m+1} + (1-\theta) u_h^m$ and $t^{m+\theta} = \theta t^{m+1} + (1-\theta)\, t^m = (m+\theta)\, k$. In matrix form, (3.6) reads

$$(k^{-1}\mathbf{K} + \theta \mathbf{A})\underline{u}^{m+1} = k^{-1}\mathbf{K}\underline{u}^m - (1-\theta)\mathbf{A}\underline{u}^m + \underline{g}^{m+\theta}, \quad m = 0, 1, ..., M-1.$$

where $\underline{u}^m$ is the coefficient vector of u_h^m with respect to a basis of V_h.

Remark 3.1. Even for the forward Euler method (i.e. for $\theta = 0$), we have to solve at each time step a linear system with the mass matrix. However, for $0 \le \rho < 1$ the spaces $H^{\rho/2}(\Omega)$ and $\tilde{H}^{\rho/2}(\Omega)$ are isomorphic and we can use discontinuous multiwavelets to obtain a diagonal mass matrix. In this case, each time step requires only one matrix-vector product with the matrix $\mathbf{A}$.

3.3 Perturbation

Previous analyses of the θ-scheme (3.5) assumed that the form $a(\cdot,\cdot)\colon V_h \times V_h \to \mathbb{R}$ can be evaluated exactly, i.e. that the corresponding stiffness matrix $\mathbf{A}$ is available. This is not the case if we compress $\mathbf{A}$, resulting in a perturbed matrix $\tilde{\mathbf{A}}$. With $\tilde{\mathbf{A}}$ we associate the form $\tilde{a}(\cdot,\cdot)$ (other perturbations e.g. due to numerical integration or domain approximation by isoparametric elements in the context of Example 2.2 can be treated in the same way). Using $\tilde{a}(\cdot,\cdot)$ in place of $a(\cdot,\cdot)$ in (3.6) gives **perturbed θ-schemes**

$$\widetilde{u}_h^0 = u_{0,h}\,, \tag{3.7a}$$

$$\Big(\frac{\widetilde{u}_h^{m+1} - \widetilde{u}_h^m}{k}, v_h\Big) + \widetilde{a}(\widetilde{u}_h^{m+\theta}, v_h) = \big(g(t^{m+\theta}), v_h\big) \tag{3.7b}$$

for $m = 0, 1, 2, \ldots, M-1$ and every $v_h \in V_h$, where again $\widetilde{u}_h^{m+\theta} := \theta \widetilde{u}_h^{m+1} + (1-\theta)\, \widetilde{u}_h^m$. In matrix form, (3.7b) reads

$$(k^{-1}\mathbf{K} + \theta \tilde{\mathbf{A}})\underline{\tilde{u}}^{m+1} = k^{-1}\mathbf{K}\underline{\tilde{u}}^m - (1-\theta)\tilde{\mathbf{A}}\underline{\tilde{u}}^m + \underline{g}^{m+\theta}, \quad m = 0, 1, ..., M-1$$

where $\underline{\tilde{u}}^m$ is the coefficient vector of $\tilde{u}_h^m$ with respect to a basis of V_h.

We shall assume for $\widetilde{a}(\cdot,\cdot)$ the following **consistency conditions**: there is $\delta < 1$ independent of h such that

$$|a(u_h, v_h) - \widetilde{a}(u_h, v_h)| \le \delta\, \|u_h\|_a\, \|v_h\|_a \quad \forall u_h, v_h \in V_h \tag{3.8}$$

and there is $C > 0$ independent of h such that

$$|a(P_h u, v_h) - \widetilde{a}(P_h u, v_h)| \le C h^{p+1-\rho/2} |\log h|^\nu \|u\|_{\mathcal{H}^{p+1}(\Omega)}\, \|v_h\|_{\tilde{H}^{\rho/2}(\Omega)} \quad \forall u \in \mathcal{H}^{p+1}(\Omega), v_h \in V_h \tag{3.9}$$

with some $\nu \ge 0$.

Condition (3.8) shows that on $V_h \times V_h$ the form $\widetilde{a}(\cdot,\cdot)$ is equivalent to $a(\cdot,\cdot)$ in the following sense:

Proposition 3.2. *For $\delta < 1$ in (3.8), we have for some constants $0 < \widetilde{\beta} \leq \widetilde{\alpha} < \infty$ independent of h*

$$\forall u_h, v_h \in V_h : \; |\widetilde{a}(u_h, v_h)| \leq \widetilde{\alpha} \, \|u_h\|_a \, \|v_h\|_a \tag{3.10}$$

and

$$\forall u_h \in V_h : \; |\widetilde{a}(u_h, u_h)| \geq \widetilde{\beta} \, \|u_h\|_a^2 \, . \tag{3.11}$$

Proof. Consider (3.11). We have for $u_h \in V_h$:

$$|\tilde{a}(u_h, u_h)| \geq |a(u_h, u_h)| - |a(u_h, u_h) - \tilde{a}(u_h, u_h)| = \|u_h\|_a^2 - |a(u_h, u_h) - \tilde{a}(u_h, u_h)|$$

and, using the definition of $\|\cdot\|_a$ and the consistency condition (3.8), we get (3.11) with $\tilde{\beta} = 1 - \delta$. The continuity (3.10) is proved in the same way. □

3.4 Wavelet Compression

In the context of Example 2.3, perturbed bilinear forms $\widetilde{a}$ are obtained by various matrix compression techniques which reduce the dense matrices $\mathbf{A}$ to sparse ones which can be manipulated in linear complexity. We illustrate this by the wavelet compression of operators of order $0 \leq \rho \leq 2$ in dimensions $d = 1, 2$; we only present here the main principles — for details and proofs, see [9, 13, 4]. All results carry over to dimensions $d > 2$ if a suitable wavelet basis is used.

3.4.1 Subspaces V_h

For $d = 1$ the domain Ω is an interval. For $d = 2$ we assume that Ω is a polygon. Let T_0 be a fixed coarse triangulation of the domain. We then define the triangulation T_l for $l > 0$ by bisection of each interval in T_{l-1} for $d = 1$, or by subdivision of a triangle in T_{l-1} in four congruent subtriangles for $d = 2$. We assume that the triangulation $\{\mathcal{T}_h\}$ is obtained in this way as T_L, for some $L > 0$ so that $h = C2^{-L}$.

For $0 \leq \rho < 1$ we define V_h as the space of piecewise polynomials of total degree $p \geq 0$ (without any continuity restriction) on the triangulation T_L.

For $1 \leq \rho \leq 2$ the space V_h is defined as the space of continuous piecewise polynomials of degree $p \geq 1$ on the triangulation with zero values on the boundary $\partial\Omega$.

In the same way we define the spaces V^l corresponding to the triangulation T_l, so that we have

$$V^0 \subset V^1 \subset \cdots \subset V^L = V_h.$$

Let $N^l = \dim V^l$ and $M^l := N^l - N^{l-1}$ so that $N = \dim V_h = N^L = C2^L$.

3.4.2 Wavelet basis

By choosing a suitable basis for V_h we will be able to represent the bilinear form $a(\cdot, \cdot)$ as a matrix where most elements are small and can be neglected, yielding the approximate bilinear form $\tilde{a}(\cdot, \cdot)$. The basis will also allow optimal preconditioning. We will use so-called biorthogonal wavelets (note that the dual wavelets described below will not be used in the computation).

We will use a hierarchical basis of functions ψ_j^l with $j = 1, \dots, M^l$ and $l = 0, 1, \dots$ with the following properties: We have $\operatorname{span}\{\psi_j^l \mid 0 \le l \le L, 1 \le j \le M^l\} = V^l$.

The function ψ_j^l has support $S_j^l := \operatorname{supp} \psi_j^l$ of diameter bounded by $C\, 2^{-l}$.

Wavelets ψ_j^l with $\bar{S}_j^l \cap \partial\Omega = \emptyset$ have vanishing moments up to order p, i.e., $(\psi_j^l, q) = 0$ for all polynomials q of total degree p or less.

The functions ψ_j^l for $l \ge l_0$ are obtained by scaling and translation of the functions $\psi_j^{l_0}$.

A function $v \in V_h$ has the representation

$$v = \sum_{l=0}^{L} \sum_{j=1}^{M^l} v_j^l \psi_j^l$$

with $v_j^l = (v, \tilde{\psi}_j^l)$ where $\tilde{\psi}_j^l$ are the so-called dual wavelets.

For $v \in V$ one obtains an infinite series

$$v = \sum_{l=0}^{\infty} \sum_{j=1}^{M^l} v_j^l \psi_j^l$$

with $v_j^l = (v, \tilde{\psi}_j^l)$ which converges in $\tilde{H}^s$ for $0 \le s \le \rho/2$.

There holds the norm equivalence

$$c_1 \|v\|_{\tilde{H}^s(\Omega)}^2 \le \sum_{l=0}^{\infty} \sum_{j=1}^{M^l} \left| v_j^l \right|^2 2^{2ls} \le c_2 \|v\|_{\tilde{H}^s(\Omega)}^2 \tag{3.12}$$

for $0 \le s \le \rho/2$ and for $\rho/2 < s \le p+1$ we have the one-sided bounds

$$\sum_{l=0}^{L} \sum_{j=1}^{M^l} \left| v_j^l \right|^2 2^{2ls} \le c_3 L^\nu \|v\|_{\mathcal{H}^s(\Omega)}^2$$

where $c_i > 0$ are independent of L, $\nu = 0$ if $s < p+1$ and $\nu = 1$ if $s = p+1$.

For $v \in V$ we can define a projection $P_h : V \to V_h$ by truncating the wavelet expansion:

$$P_h v := \sum_{l=0}^{L} \sum_{j=1}^{M^l} v_j^l \psi_j^l \tag{3.13}$$

This projection satisfies the approximation property (3.3).

3.4.3 Examples for wavelets

In the case $0 \le \rho < 1$ a multiwavelet basis as in [11] can be used: Let $\{p_k\}$ be a basis for polynomials of total degree p or less. Then the functions ψ_j^0 are the functions which are on one element equal to a function p_k, and zero elsewhere. For $l > 1$ we choose for ψ_j^l functions in V^l which are nonzero on one element of T_{l-1}, and which are orthogonal on all polynomials of total degree p or less.

In all cases $0 \le \rho \le 2$ so-called prewavelets can be used: These are functions in V^l with small support which are orthogonal on V^{l-1}.

Another possiblity are so-called biorthogonal wavelets which need not be orthogonal on V^{l-1}. For piecewise linears the functions ψ_j^l in the interior of the interval have values $0,\ldots,0,-1,2,-1,0,\ldots,0$. In the case of Neumann boundary conditions the wavelet at the left boundary has values $-2,2,1,0,\ldots,0$; in the case of Dirichlet conditions the values are $0,2,-1,0,\ldots,0$ (and similarly at the right boundary). Note that the boundary wavelets have fewer vanishing moments in general.

In dimension $d=2$ the construction of piecewise linear prewavelets on arbitrary polygons is described e.g. in [5, 15].

3.4.4 Matrix compression

The bilinear form a on $V_h \times V_h$ corresponds to a matrix $\mathbf{A}$ with elements $A_{(l,j),(l',j')} = a(\psi_j^l, \psi_{j'}^{l'})$.

We assumed that the kernel of the operator satisfies the estimates (2.18). This implies a decay of the matrix elements with increasing distance of their supports.

We define the compressed matrix $\tilde{\mathbf{A}}$ and the corresponding bilinear form $\tilde{a}$ by replacing certain small matrix elements in $\mathbf{A}$ with zero:

$$\tilde{A}_{(j,l),(j',l')} := \begin{cases} A_{(j,l),(j',l')} & \text{if } \operatorname{dist}(S_j^l, S_{j'}^{l'}) \leq \delta_{l,l'} \text{ or } S_j^l \cap \partial\Omega \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

Here the truncation parameters $\delta_{l,l'}$ are given by

$$\delta_{l,l'} := c \max\{2^{-L+\hat{\alpha}(2L-l-l')}, 2^{-l}, 2^{-l'}\} \tag{3.15}$$

with some parameters $c > 0$ and $\hat{\alpha} > 0$. The consistency conditions (3.8), (3.9) can be satisfied (see e.g. [9, 10, 11, 13]):

Proposition 3.3. *If c in (3.15) is chosen sufficiently large then for all $L > 0$ condition (3.8) holds. If additionally*

$$\hat{\alpha} \geq \frac{2p+2}{2p+2+\rho} \tag{3.16}$$

holds then condition (3.9) holds with $\nu = \frac{3}{2}$ if equality holds in (3.16), and $\nu = \frac{1}{2}$ otherwise.

The matrix compression (3.14) reduces the number of nonzero elements from N^2 to N times a logarithmic term [9, 10, 11, 13]:

Proposition 3.4. *The compressed matrix $\tilde{\mathbf{A}}$ has $O(N \log N)$ nonzero elements if $\hat{\alpha} < 1$, and $O(N(\log N)^2)$ nonzero elements if $\hat{\alpha} = 1$.*

In particular, for operators of order $\rho > 0$ we can choose $\hat{\alpha}$ such that $\nu = \frac{1}{2}$ in (3.9) and the number of nonzero elements in $\tilde{\mathbf{A}}$ is $O(N \log N)$. In the case of order $\rho = 0$ we have to choose $\hat{\alpha} = 1$ implying $\nu = \frac{3}{2}$ in (3.9) and the number of nonzero elements in $\tilde{\mathbf{A}}$ is $O(N(\log N)^2)$.

4 Stability

The stability of the θ-scheme is well-known in the context of Example 2.2, i.e., if the spatial operator is elliptic and of second order. We investigate here general operators A of order $\rho \geq 0$ that are elliptic in the sense that (2.7), (2.8) hold in $V = \widetilde{H}^{\rho/2}(\Omega)$. We prove an $L^2(J,V)$ stability estimate for the approximate solutions obtained from the θ-scheme with wavelet-compressed space operator.

In the analysis, we will use for $f \in V_h^*$ the following notation:

$$\|f\|_* := \sup_{v_h \in V_h} \frac{(f, v_h)}{\|v_h\|_a} . \tag{4.1}$$

We will also need λ_A defined by

$$\lambda_A := \sup_{v_h \in V_h} \frac{\|v_h\|^2}{\|v_h\|_*^2}$$

We first address the stability of the θ-scheme with exact bilinear form $a(\cdot, \cdot)$. In the case $\frac{1}{2} \leq \theta \leq 1$, the θ-scheme is stable for any time step $k > 0$, whereas in the case $0 \leq \theta < \frac{1}{2}$ the time step k must be sufficiently small.

Proposition 4.1. *In the case of* $\frac{1}{2} \leq \theta \leq 1$ *assume that*

$$0 < C_1 < 2, \qquad C_2 \geq \frac{1}{2 - C_1} \tag{4.2}$$

and in the case of $0 \leq \theta < \frac{1}{2}$ *assume that*

$$\sigma := k(1 - 2\theta)\lambda_A < 2 \tag{4.3}$$

$$0 < C_1 < 2 - \sigma, \qquad C_2 \geq \frac{1 + (4 - C_1)\sigma}{2 - \sigma - C_1}. \tag{4.4}$$

Then the sequence $\{u_h^m\}_{m=0}^M$ *of solutions of the θ-scheme* (3.5) *satisfies the stability estimate*

$$\|u_h^M\|^2 + C_1 k \sum_{m=0}^{M-1} \|u_h^{m+\theta}\|_a^2 \leq \|u_h^0\|^2 + C_2 k \sum_{m=0}^{M-1} \|g^{m+\theta}\|_*^2. \tag{4.5}$$

Proof. Let

$$X^m := \|u_h^m\|^2 - \left\|u_h^{m+1}\right\|^2 + C_2 k \left\|g^{m+\theta}\right\|_*^2 - C_1 k \left\|u_h^{m+\theta}\right\|_a^2$$

We want to show that $X^m \geq 0$. Then adding these inequalities for $m = 0, \ldots, M-1$ will obviously give (4.5).

Let $w := u_h^{m+1} - u_h^m$, then $u_h^{m+\theta} = (u_h^m + u_h^{m+1})/2 + (\theta - \frac{1}{2})w$ and

$$\left\|u_h^{m+1}\right\|^2 - \|u_h^m\|^2 = (u_h^{m+1} - u_h^m, u_h^{m+1} + u_h^m) = (w, 2u_h^{m+\theta} - (2\theta - 1)w).$$

By the definition of the θ-scheme we have

$$\begin{aligned}(w, u_h^{m+\theta}) &= k(-Au_h^{m+\theta} + g^{m+\theta}, u_h^{m+\theta}) = k\Big[-\left\|u_h^{m+\theta}\right\|_a^2 + (g^{m+\theta}, u_h^{m+\theta})\Big] \\ &\leq k\Big[-\left\|u_h^{m+\theta}\right\|_a^2 + \|g^{m+\theta}\|_* \left\|u_h^{m+\theta}\right\|_a\Big]\end{aligned}$$

This gives

$$X^m \geq (2\theta - 1)\|w\|^2 + k\Big[(2 - C_1)\left\|u_h^{m+\theta}\right\|_a^2 - 2\left\|g^{m+\theta}\right\|_* \left\|u_h^{m+\theta}\right\|_a + C_2 \left\|g^{m+\theta}\right\|_*^2\Big]$$

In the case of $\frac{1}{2} \leq \theta \leq 1$ we now obtain $X^m \geq 0$ if the conditions (4.2) are satisfied.

In the case $0 \leq \theta < \frac{1}{2}$ we have by the definition of the θ-scheme that $(w, v_h) = k(-Au_h^{m+\theta} + g^{m+\theta}, v_h)$ yielding

$$\|w\| \leq \lambda_A^{1/2} \|w\|_* \leq \lambda_A^{1/2} k\left(\left\|Au_h^{m+\theta}\right\|_* + \left\|g^{m+\theta}\right\|_*\right) = \lambda_A^{1/2} k\left(\left\|u_h^{m+\theta}\right\|_a + \left\|g^{m+\theta}\right\|_*\right)$$

since $(Au_h^{m+\theta}, v_h) \leq \left\|u_h^{m+\theta}\right\|_a \|v_h\|_a$ gives $\left\|Au_h^{m+\theta}\right\|_* \leq \left\|u_h^{m+\theta}\right\|_a$ and choosing $v_h := u_h^{m+\theta}$ gives $\left\|Au_h^{m+\theta}\right\|_* \geq \left\|u_h^{m+\theta}\right\|_a$. Hence

$$k^{-1} X^m \geq (2 - C_1 - \sigma) \left\|u_h^{m+\theta}\right\|_a^2 - 2(1+\sigma) \left\|g^{m+\theta}\right\|_* \left\|u_h^{m+\theta}\right\|_a + (C_2 - \sigma) \left\|g^{m+\theta}\right\|_*^2 .$$

Therefore we have $X^m \geq 0$ if conditions (4.3) hold. $\square$

Remark 4.2. The a-priori estimate (4.5) is, in a sense, the discrete analogue of the a-priori estimate (2.14). Note, however, that $\|g^{m+\theta}\|_*$ is not identical to $\|g^{m+\theta}\|_{V^*}$—in fact for $g \in V^*$, we have by $V_h \subset V$ that

$$\|g^{m+\theta}\|_* \leq \|g^{m+\theta}\|_{V^*} .$$

Consider now the sequence $\{\tilde{u}_h^m\}_{m=0}^M$ of solutions to the perturbed θ-scheme (3.7a), (3.7b). We analogously define for $v_h \in V_h$ and $f \in V_h^*$

$$\|v_h\|_{\tilde{a}} := \tilde{a}(v_h, v_h), \qquad \|f\|_{\tilde{*}} := \sup_{v_h \in V_h} \frac{(f, v_h)}{\|v_h\|_{\tilde{a}}}, \qquad \lambda_{\tilde{A}} := \sup_{v_h \in V_h} \frac{\|v_h\|^2}{\|v_h\|_{\tilde{*}}^2} \tag{4.6}$$

Due to the norm equivalence in Proposition 3.2, we obtain in the same way as in Proposition 4.1 with $\tilde{a}(\cdot,\cdot)$ in place of $a(\cdot,\cdot)$

Proposition 4.3. *Assume that* (3.8) *holds with* $\delta < 1$. *In the case of* $\frac{1}{2} \leq \theta \leq 1$ *assume that* (4.2) *holds. In the case of* $0 \leq \theta < \frac{1}{2}$ *assume that*

$$\sigma := k(1 - 2\theta)\lambda_{\tilde{A}} < 2 \tag{4.7}$$

and that (4.4) *holds.*
Then the sequence $\{\tilde{u}_h^m\}_{m=0}^M$ *of solutions of the perturbed* θ*-scheme* (3.7a), (3.7b) *satisfies the stability estimate*

$$\|\tilde{u}_h^M\|^2 + C_1 k \sum_{m=0}^{M-1} \|\tilde{u}_h^{m+\theta}\|_{\tilde{a}}^2 \leq \|\tilde{u}_h^0\|^2 + C_2 k \sum_{m=0}^{M-1} \|g^{m+\theta}\|_{\tilde{*}}^2. \tag{4.8}$$

Remark 4.4. By the inverse estimate (3.4) and the norm equivalence (2.14) we have for $w_h \in V_h$

$$\|w_h\|_a \leq C \|w_h\|_{\rho/2} \leq C' h^{-\rho/2} \|w_h\|$$

and therefore for $v_h \in V_h$

$$\|v_h\|_* = \sup_{w_h \in V_h} \frac{(v_h, w_h)}{\|w_h\|_a} \geq C h^{\rho/2} \sup_{w_h \in V_h} \frac{(v_h, w_h)}{\|w_h\|} = C h^{\rho/2} \|v_h\| \tag{4.9}$$

$$\lambda_A^{1/2} = \sup_{v_h \in V_h} \frac{\|v_h\|}{\|v_h\|_*} \leq C h^{-\rho/2}. \tag{4.10}$$

Hence there exists a positive constant C_* independent of h and θ such that the time-step restriction

$$k \leq C_* \frac{h^\rho}{1-2\theta} \tag{4.11}$$

is sufficient for stability (4.3). For $\rho = 2$ and $\theta < \frac{1}{2}$ (e.g., forward Euler and the heat equation) this reduces to the well-known time-step restriction $k \leq C_\theta h^2$ for explicit schemes. For smaller values of ρ the restriction is less severe, and in the limiting case $\rho = 0$ condition (4.11) gives $k \leq C_*/(1-2\theta)$ with a bound independent of h.

For the perturbed scheme (3.7) we can proceed in the same way and obtain using Proposition 3.2 that (4.11) is a sufficient condition for (4.7) (with a different value of C_*).

Remark 4.5. As θ tends to $\frac{1}{2}$ from below the bound on k in the stability condition (4.3) tends to infinity, and for $\theta \geq \frac{1}{2}$ the stability holds with $\sigma = 0$ and C_1, C_2 as in (4.4) for *all* values of k.

5 Convergence

Based on the stability results obtained in Section 4 and the consistency (3.8), (3.9) of the compressed form $\widetilde{a}(\cdot,\cdot)$, we shall now obtain optimal convergence estimates of the compressed θ-scheme (sufficient regularity of the exact solution $u(x,t)$ in space and time provided). Throughout this section, we shall set

$$u^m = u(t^m) \in V. \tag{5.1}$$

We will estimate the error

$$\widetilde{e}_h^m := u^m - \widetilde{u}_h^m \,. \tag{5.2}$$

To this end, we split $\widetilde{e}_h^m$ as follows:

$$\widetilde{e}_h^m = \underbrace{(u^m - P_h u^m)}_{\eta^m} + \underbrace{(P_h u^m - \widetilde{u}_h^m)}_{\xi_h^m} = \eta^m + \xi_h^m \tag{5.3}$$

where $P_h : V \to V_h$ is the quasi-interpolant in (3.13) (realized as a truncated wavelet expansion, see Section 3.4 or [9, 3] for details).

As η^m is a best approximation error, we focus now on $\xi_h^m \in V_h$.

Lemma 5.1. *The $\{\xi_h^m\}_m$ are solutions of the θ-scheme*

$$\xi_h^0 = P_h u_0 - \widetilde{u}_h^0 \,,$$

for $m = 0, 1, \ldots, M-1$ and every $v_h \in V_h$:

$$k^{-1}(\xi_h^{m+1} - \xi_h^m, v_h) + \widetilde{a}\,(\theta \xi_h^{m+1} + (1-\theta)\,\xi_h^m, v_h) = (r^m, v_h) \tag{5.4}$$

where the weak residuals r^m: $V_h \to \mathbb{R}$ are given by

$$r^m = r_1^m + r_2^m + r_3^m + r_4^m \tag{5.5}$$

with

$$
\begin{aligned}
(r_1^m, v_h) &:= \Big(\frac{u^{m+1}-u^m}{k} - \dot u^{m+\theta}, v_h\Big)\,,\\
(r_2^m, v_h) &:= \Big(\frac{P_h u^{m+1}-P_h u^m}{k} - \frac{u^{m+1}-u^m}{k}\,,\ v_h\Big)\,,\\
(r_3^m, v_h) &:= \widetilde a\,(P_h u^{m+\theta}, v_h) - a\,(P_h u^{m+\theta}, v_h)\,,\\
(r_4^m, v_h) &:= a\,(P_h u^{m+\theta} - u^{m+\theta}, v_h)\,.
\end{aligned}
$$

Proof. We note that (2.11) and $u \in C(\overline{J}, H)$ imply

$$
(\dot u^{m+\theta}, v) + a(u^{m+\theta}, v) = (g^{m+\theta}, v) \quad \forall v \in V\,. \tag{5.6}
$$

Since $V_h \subset V$, we get for every $v_h \in V_h$

$$
\begin{aligned}
& k^{-1}(\xi_h^{m+1} - \xi_h^m, v_h) + \widetilde a(\theta \xi_h^{m+1} + (1-\theta)\,\xi_h^m, v_h)\\
= \ & \Big(\frac{(P_h u^{m+1} - \widetilde u_h^{m+1}) - (P_h u^m - \widetilde u_h^m)}{k}\,,\ v_h\Big) + \widetilde a\,(P_h u^{m+\theta}, v_h) - \widetilde a\,(\widetilde u_h^{m+\theta}, v_h)\\
= \ & \Big(\frac{P_h u^{m+1} - P_h u^m}{k}\,,\ v_h\Big) + \widetilde a(P_h u^{m+\theta}, v_h) - \Big\{\Big(\frac{\widetilde u_h^{m+1} - \widetilde u_h^m}{k}\,,\ v_h\Big) - \widetilde a\,(\widetilde u_h^{m+\theta}, v_h)\Big\}\\
\overset{(3.7)}{=} \ & \Big(\frac{P_h u^{m+1} - P_h u^m}{k}\,,\ v_h\Big) + \widetilde a\,(P_h u^{m+\theta}, v_h) - (g^{m+\theta}, v_h)\\
\overset{(5.6)}{=} \ & \Big(\frac{P_h u^{m+1} - P_h u^m}{k} - \dot u^{m+\theta}, v_h\Big) + \widetilde a\,(P_h u^{m+\theta}, v_h) - a\,(u^{m+\theta}, v_h) =: (r, v_h)\,.
\end{aligned}
$$

The representation (5.5) of (r, v_h) is now evident. ☐

Lemma 5.1 implies together with the stability result Proposition 4.3 the following estimate for the ξ_h^m:

Corollary 5.2. *Under the assumptions of Proposition* 4.3*, we have*

$$
\|\xi_h^M\|^2 + C_1 k \sum_{m=0}^{M-1} \big\|\xi_h^{m+\theta}\big\|_{\widetilde a}^2 \le \|\xi_h^0\|^2 + C_2 k \sum_{m=0}^{M-1} \big\|r^m\big\|_{\widetilde *}^2 \tag{5.7}
$$

Based on (5.5), we must estimate the $\|r_j^m\|_{\widetilde *}$, $j = 1, \ldots, 4$.

Estimate of r_1^m: is based on Taylor expansion in t. Noting that for any $v_h \in V_h$

$$
|(r_1^m, v_h)| \le \|k^{-1}(u^{m+1} - u^m) - \dot u^{m+\theta}\|_* \,\|v_h\|_a\,,
$$

and

$$
k^{-1}(u^{m+1} - u^m) - \dot u^{m+\theta} = \frac{1}{k} \int_{t_m}^{t_{m+1}} (s - (1-\theta)t_{m+1} - \theta\, t_m)\,\ddot u\, ds\,,
$$

we get

$$\begin{aligned}
\|k^{-1}(u^{m+1} - u^m) - \dot u^{m+\theta}\|_* &\le k^{-1} \int\limits_{t_m}^{t_{m+1}} |s - (1-\theta)\, t_{m+1} - \theta\, t_m|\; \|\ddot u\|_*\, ds \\
&\le C_\theta \int\limits_{t_m}^{t_{m+1}} \|\ddot u(s)\|_*\, ds \qquad (5.8)\\
&\le C_\theta\, k^{+\frac12} \Big(\int\limits_{t_m}^{t_{m+1}} \|\ddot u(s)\|_*^2\, ds \Big)^{\frac12} .
\end{aligned}$$

If $\theta = \frac12$, an integration by parts gives

$$k^{-1}(u^{m+1} - u^m) - \dot u^{m+\theta} = \frac{1}{2k} \int\limits_{t_m}^{t_{m+1}} (t_{m+1} - s)(t_m - s)\; \dddot u\,(s) ds$$

and it follows that

$$\|k^{-1}(u^{m+1} - u^m) - \dot u^{m+\theta}\|_* = C\, k^{\frac32} \Big(\int\limits_{t_m}^{t_{m+1}} \|\; \dddot u\,(s)\|_*^2\, ds \Big)^{\frac12} .$$

Estimate of r_2^m: here

$$\begin{aligned}
|(r_2^m, v_h)| \;&\le\; C\, \|k^{-1} \left[(u^{m+1} - u^m) - P_h(u^{m+1} - u^m)\right]\|_*\, \|v_h\|_{\widetilde a} \\
&=\; C\, k^{-1} \left\| (I - P_h) \int\limits_{t_m}^{t_{m+1}} \dot u(s)\, ds \right\|_* \|v_h\|_{\widetilde a} \qquad (5.9)\\
&\le\; C\, k^{-1} \int\limits_{t_m}^{t_{m+1}} \|(I - P_h)\, \dot u\|_*\, ds\, \|v_h\|_{\widetilde a}
\end{aligned}$$

where we may now use the approximation property (3.3) of P_h pointwise in t.

Estimate of r_3^m: here we use the consistency (3.9)

$$|(r_3^m, v_h)| \le C\, h^{p+1-\rho/2} |\log h|^\nu \|u^{m+\theta}\|_{\mathcal{H}^{p+1}(\Omega)} \|v_h\|_{\tilde H^{\rho/2}(\Omega)} . \qquad (5.10)$$

By (2.14) and (3.11), we get $\|v_h\|_{\tilde H^{\rho/2}} \le C\, \|v_h\|_{\widetilde a}$ and hence a bound on $\|r_3^m\|_{\widetilde *}$.

Estimate on r_4^m: Using (3.9) with $s = 0$ and (3.3) gives

$$|(r_4^m, v_h)| \le C\, \|u^{m+\theta} - P_h u^{m+\theta}\|_a\, \|v_h\|_{\widetilde a}$$

and with the approximation property (3.3) we find

$$|(r_4^m, v_h)| \le C\, h^{p+1-\rho/2} \|u^{m+\theta}\|_{\mathcal{H}^{p+1}(\Omega)}\, \|v_h\|_{\widetilde a} . \qquad (5.11)$$

Collecting the bounds (5.8)–(5.11) gives

Lemma 5.3. *Assume that (3.8), (3.9) hold. If $u(x,t)$ is sufficiently smooth in $\overline{J}\times\overline{\Omega}$, we have for r^m given by* (5.5)

$$\|r^m\|_{\widetilde{*}} \le C \begin{cases} k^{\frac12}\Big(\int\limits_{t_m}^{t_{m+1}} \|\ddot{u}(s)\|_*^2 ds\Big)^{\frac12} & \textit{for all } \theta\in[0,1] \\ k^{\frac32}\Big(\int\limits_{t_m}^{t_{m+1}} \|\dddot{u}(s)\|_*^2 ds\Big)^{\frac12} & \textit{for } \theta=\frac12 \end{cases} \tag{5.12}$$
$$+ Ck^{-\frac12}h^{p+1-\rho/2}\Big(\int\limits_{t_m}^{t_{m+1}} \|\dot{u}\|^2_{\mathcal{H}^{p+1-\rho/2}(\Omega)} ds\Big)^{\frac12}$$
$$+ Ch^{p+1-\rho/2}|\log h|^{\nu}\|u^{m+\theta}\|_{\mathcal{H}^{p+1}(\Omega)}\,.$$

Theorem 5.4. *Assume that the consistency conditions (3.8), (3.9) hold. For $\theta\in[0,\frac12)$ assume (4.7). Assume further that the approximation $u_{0,h}\in V_h$ of the initial data u^0 is quasioptimal in $L^2(\Omega)$. Then holds the following error estimate for the perturbed θ-scheme with $\theta\in[0,1]$*

$$\big\|u^M-\widetilde{u}_h^M\big\|^2 + k\sum_{m=0}^{M-1}\|u^{m+\theta}-\widetilde{u}_h^{m+\theta}\|_a^2 \le Ch^{2(p+1-\rho/2)}|\log h|^{2\nu}\max_{0\le t\le T}\|u(t)\|^2_{\mathcal{H}^{p+1}(\Omega)}$$
$$+ C\begin{cases} k^2\int\limits_0^T \|\ddot{u}(s)\|_*^2 ds & \textit{for all } \theta\in[0,1] \\ k^4\int\limits_0^T \|\dddot{u}(s)\|_*^2 ds & \textit{for } \theta=\frac12 \end{cases} \tag{5.13}$$
$$+ Ch^{2(p+1-\rho/2)}\int\limits_0^T \|\dot{u}(s)\|^2_{\mathcal{H}^{p+1-\rho/2}(\Omega)} ds\,.$$

Proof. Based on (5.3), we have for every $M\ge1$

$$\big\|\widetilde{e}_h^M\big\|^2 + k\sum_{m=0}^{M-1}\|\widetilde{e}_h^{m+\theta}\|_a^2 \le 2\left\{\big\|\eta^M\big\|^2 + k\sum_{m=0}^{M-1}\|\eta^{m+\theta}\|_a^2\right\} + 2\left\{\big\|\xi_h^M\big\|^2 + k\sum_{m=0}^{M-1}\|\xi_h^{m+\theta}\|_a^2\right\}.$$

The first term can be estimated with the approximation property (3.3). The second term is treated using (3.11) and (4.8). We get

$$\begin{aligned}
\big\|\xi_h^M\big\|^2 + k\sum_{m=0}^{M-1}\|\xi_h^{m+\theta}\|_a^2 &\le \big\|\xi_h^M\big\|^2 + k\widetilde{\beta}^{-1}\sum_{m=0}^{M-1}\Big\|\xi_h^{m+\theta}\Big\|_{\widetilde{a}}^2 \\
&\le \max\{1,1/(\widetilde{\beta}C_1)\}\left\{\big\|\xi_h^M\big\|^2 + C_1k\sum_{m=0}^{M-1}\Big\|\xi_h^{m+\theta}\Big\|_{\widetilde{a}}^2\right\} \\
&\le \max\{1,1/(\widetilde{\beta}C_1)\}\left\{\big\|\xi_h^0\big\|^2 + C_2k\sum_{m=0}^{M-1}\Big\|r_h^{m+\theta}\Big\|_{\widetilde{*}}^2\right\}
\end{aligned}$$

Using now the bound (5.12) for $\Big\|r_h^{m+\theta}\Big\|_{\widetilde{*}}$, the quasioptimality of $u_{0,h}$ and the approximation property (3.3) with $s=0$ to estimate $\big\|\xi_h^0\big\|$ gives the assertion. □

6 Approximate Solution of Linear Equations and Complexity

In order to compute the approximate solution $\tilde{u}_h^m$ in (3.7) for $m = 1, \ldots, M$ we proceed as follows:

We first compute the mass matrix $\mathbf{K}$ in the wavelet basis with elements $K_{(l,j),(l',j')}$ where $O(N \log N)$ elements are nonzero. Note that for discontinuous multiwavelets (which can be used for $0 \le \rho < 1$) the mass matrix is diagonal.

Then we compute the compressed stiffness matrix $\tilde{\mathbf{A}}$ where $O(N(\log N)^r)$ elements are nonzero and $r = 1$ if $\rho \in (0, 2]$, $r = 2$ if $\rho = 0$, see Proposition 3.4. If explicit antiderivatives of the kernel function are available (as is often the case), the total cost for computing the stiffness matrix $\tilde{\mathbf{A}}$ is $O(N(\log N)^r)$ operations. In other cases quadrature as described in [10] can be used. This preserves the consistency conditions (3.8),(3.9) and the total cost of computing $\tilde{\mathbf{A}}$ is $O(N(\log N)^{r+d})$ for $d = 1, 2$.

For each time step we have to solve (3.7b): We have to find $\tilde{w}_h^m := \tilde{u}_h^{m+1} - \tilde{u}_h^m \in V_h$ satisfying

$$k^{-1}(\tilde{w}_h^m, v_h) + \theta \tilde{a}(\tilde{w}_h^m, v_h) = (g^{m+\theta}, v_h) - \tilde{a}(\tilde{u}_h^m, v_h) \quad \forall v_h \in V_h \tag{6.1}$$

and then update $\tilde{u}_h^{m+1} := \tilde{u}_h^m + \tilde{w}_h^m$. Let $\tilde{w}^m \in \mathbb{R}^N$ denote the coefficient vectors of $\tilde{w}_h^m$ with respect to the wavelet basis, and $\mathbf{K}, \tilde{\mathbf{A}} \in \mathbb{R}^{N \times N}$ the mass and stiffness matrices corresponding to $(\cdot, \cdot)$ and $\tilde{a}(\cdot, \cdot)$ in this basis. Then we obtain for $\tilde{w}^m$ a linear system $\mathbf{B}\tilde{w}^m = \tilde{b}^m$ with the matrix $\mathbf{B} = k^{-1}\mathbf{K} + \theta\tilde{\mathbf{A}}$ and a known right-hand side vector $\tilde{b}^m$.

For a standard finite element basis, the matrix $\mathbf{B}$ has a condition number of order $h^{-\rho}$ for small h and fixed k. For the matrix $\mathbf{B}$ in the wavelet basis we can achieve a uniformly bounded condition number if we scale the rows and columns of $\mathbf{B}$ as follows: let $\mu_l := (k^{-1} + \theta 2^{l\rho})^{1/2}$ and let $\hat{B}_{(l,j),(l',j')} := \mu_l^{-1} \mu_{l'}^{-1} B_{(l,j),(l',j')}$. A similar scaling was proposed in [4]. Let in what follows $\|\cdot\|$ denote the 2-norm of a vector, or the 2-norm of a matrix.

Lemma 6.1. *For the linear system $\hat{\mathbf{B}}x = b$ let x_i for $i \in \mathbb{N}$ denote the iterates obtained by the restarted GMRES(m_0) method with initial guess x_0. Then there holds*

$$\|x - x_j\| \le C q^j \|x - x_0\| \tag{6.2}$$

where C and $q < 1$ are independent of L, k, θ.

Proof. Throughout the proof, C_i will denote generic positive constants independent of h,k,m, unrelated to C_i above. Let $\mathbf{D}$ denote the diagonal matrix with entries $D_{(l,j),(l,j)} = 2^{l\rho/2}$. Because of the norm equivalence (3.12) we have for all $x, y \in \mathbb{R}^N$

$$C_1 \|x\|^2 \le x^T \mathbf{K} x, \qquad x^T \mathbf{K} y \le C_2 \|x\| \|y\|$$

Using the consistency conditions (3.8) of the wavelet truncation and (2.7), (2.8) we obtain

$$C_3 \|\mathbf{D}x\|^2 \le x^T \tilde{\mathbf{A}} x, \qquad x^T \tilde{\mathbf{A}} y \le C_4 \|\mathbf{D}x\| \|\mathbf{D}y\|$$

The constants $C_j > 0$ are independent of L. Therefore $\mathbf{B} = k^{-1}\mathbf{K} + \theta\tilde{\mathbf{A}}$ satisfies with $C_5 := \min\{C_1, C_3\}$ and $C_6 := \max\{C_2, C_4\}$

$$C_5\, x^T (k^{-1}\mathbf{I} + \theta \mathbf{D}^2) x \le x^T \mathbf{B} x \tag{6.3}$$

$$x^T \mathbf{B} y \le C_6 \left[k^{-1} \|x\| \|y\| + \theta \|\mathbf{D}x\| \|\mathbf{D}y\| \right] \le C_6 \left[x^T (k^{-1}\mathbf{I} + \theta\mathbf{D}^2) x \right]^{1/2} \left[y^T (k^{-1} + \theta \mathbf{D}^2) y \right]^{1/2} \tag{6.4}$$

using the Cauchy-Schwarz inequality for the last estimate. Hence scaling with the diagonal matrix $\mathbf{S} := (k^{-1}\mathbf{I} + \theta\mathbf{D}^2)^{1/2}$ yields with $\hat{\mathbf{B}} = \mathbf{S}^{-1}\mathbf{B}\mathbf{S}^{-1}$ and $\hat{x} := \mathbf{S}x$, $\hat{y} := \mathbf{S}y$ that

$$C_5 \|\hat{x}\|^2 \leq \hat{x}^T \hat{\mathbf{B}} \hat{x}, \qquad \hat{x}^T \hat{\mathbf{B}} \hat{y} \leq C_6 \|\hat{x}\| \|\hat{y}\| \tag{6.5}$$

for all $\hat{x}, \hat{y} \in \mathbb{R}^N$ and therefore

$$\lambda_{\min}\big((\hat{\mathbf{B}} + \hat{\mathbf{B}}^T)/2\big) \geq C_5, \qquad \left\|\hat{\mathbf{B}}\right\| \leq C_6$$

According to [6] the non-restarted GMRES method for the matrix $\hat{\mathbf{B}}$ therefore satisfies for the iterates x_m and their residuals $r_m := b - \hat{\mathbf{B}} x_m$

$$\|r_m\| \leq \Big(1 - \frac{C_5^2}{C_6^2}\Big)^{m/2} \|r_0\|.$$

Because of $C_5 \|x_m - x\|^2 \leq (x_m - x)^T \hat{\mathbf{B}} (x_m - x) \leq C_6 \|x_m - x\| \|r_m\|$ a corresponding estimate holds for the errors $\|x_m - x\|$. ☐

Remark 6.2. *If the operator A is symmetric we can also use the conjugate gradient method for the symmetric matrix $\hat{\mathbf{B}}$. This will in general give the bound (6.2) with a smaller constant q than the GMRES method.*

Note that for a function $v_h \in V_h$ with coefficient vector v and scaled coefficient vector $\hat{v} = \mathbf{S}v$ we have from (6.5) that with $b(u,v) := k^{-1}(u,v) + \theta\tilde{a}(u,v)$ and $\|v\|_b^2 := b(v,v)$

$$\|\hat{v}\|^2 \sim \hat{v}^T \hat{\mathbf{B}} \hat{v} = \|v_h\|_b^2.$$

A functional $f_h \in V_h^*$ corresponds to a coefficient vector f so that $(f_h, v_h) = f^T v$, and a scaled vector $\hat{f} = \mathbf{S}^{-1} f$ so that $(f_h, v_h) = \hat{f}^T \hat{v}$. Assume that we solve a linear system $\hat{B}\hat{v}_* = \hat{f}$ using n_G steps of GMRES(m_0), starting with initial guess 0, yielding an approximation $\hat{v}$. We then have

$$\|v_{h,*} - v_h\|_b \leq C q^{n_G} \|v_{h,*}\|_b$$

and for the residuals $\rho_h \in V_h^*$ defined by $(\rho_h, w_h) = (f, w_h) - b(v_h, w_h)$ it holds that

$$\|\rho_h\|_{b,*} \leq C q^{n_G} \|f_h\|_{b,*}$$

where for $g_h \in V_h^*$ with $\hat{\mathbf{B}}_s := (\hat{\mathbf{B}} + \hat{\mathbf{B}}^T)/2$

$$\|g_h\|_{b,*} := \sup_{w_h \in V_h} \frac{(g_h, w_h)}{\|w_h\|_b} = (\hat{g}^T \hat{\mathbf{B}}_s^{-1} \hat{g})^{1/2} \sim \|\hat{g}\|.$$

We have with the inverse inequality

$$(c_1 k^{-1} h^\rho + \theta)\tilde{a}(v_h, v_h) \leq b(v_h, v_h) \leq (c_2 k^{-1} + \theta)\tilde{a}(v_h, v_h)$$

implying

$$(c_2 k^{-1} + \theta)^{-1/2} \|f_h\|_{\tilde{*}} \leq \|f_h\|_{b,*} \leq (c_1 k^{-1} h^\rho + \theta)^{-1/2} \|f_h\|_{\tilde{*}}$$

and

$$\|v_{h,*} - v_h\|_{\tilde{a}} \leq C\gamma^{1/2} q^{n_G} \|v_{h,*}\|_{\tilde{a}}, \tag{6.6}$$

$$\|\rho_h\|_{\tilde{*}} \leq C\gamma^{1/2} q^{n_G} \|f_h\|_{\tilde{*}}. \tag{6.7}$$

where

$$\gamma := \frac{c_2 k^{-1} + \theta}{c_1 k^{-1} h^\rho + \theta}.$$

We now define the **perturbed θ-scheme with GMRES approximation** as follows: Pick a value $m_0 \geq 1$ for the restart number, e.g., $m_0 = 1$, and a value n_G for the number of iterations. Let $\check{u}_h^0 := u_{0,h}$. At each time step we want to want to find an approximation of $w_{h,*}^m$ satisfying

$$b(w_{h,*}^m, v_h) = (g^{m+\theta}, v_h) - \tilde{a}(\check{u}_h^m, v_h) \qquad \text{for all } v_h \in V_h$$

which corresponds to a scaled linear system $\hat{\mathbf{B}}\hat{w}_*^m = \hat{b}^m$. We solve this system approximately with n_G steps of GMRES(m_0), using zero as initial guess, yielding an approximation $\hat{w}^m$ of the exact solution $\hat{w}_*^m$. We then let $\check{u}_h^{m+1} := \check{u}_h^m + w_h^m$ where $w_h^m \in V_h$ is the function corresponding to the scaled vector $\hat{w}^m$. Then we have

Theorem 6.3. *Assume that the consistency conditions (3.8), (3.9) hold. For $\theta \in [0, \frac{1}{2})$ assume (4.3). Then the solution $\check{u}_h^m$ of the perturbed θ-scheme with GMRES approximation satisfies the same error bound as $\tilde{u}_h^m$ in (5.13) if $n_G \geq C\,|\log h|$.*

Proof. Let $\tilde{u}_h^m$ denote the solution of (3.7) (with all linear systems solved exactly), and let $\check{u}_h^m$ denote the corresponding solution where the linear system (6.1) for each time step is solved with n_G GMRES(m_0) steps, using zero as initial guess. Let $\rho_h^m \in V_h^*$ denote the residual of the approximate GMRES solution w_h^m: For all $v_h \in V_h$

$$(\rho_h^m, v_h) = b(w_h^m, v_h) - (g^{m+\theta}, v_h) + \tilde{a}(\check{u}_h^m, v_h) = k^{-1}(\check{u}_h^{m+1} - \check{u}_h^m, v_h) + \tilde{a}(\check{u}_h^{m+\theta}, v_h) - (g^{m+\theta}, v_h)$$

Then the difference $\zeta_h^m := \check{u}_h^m - \tilde{u}_h^m$ satisfies $\zeta_h^0 = 0$ and a θ-scheme of the same form as (3.7b)

$$k^{-1}(\zeta_h^{m+1} - \zeta_h^m, v_h) + \tilde{a}(\zeta_h^{m+\theta}, v_h) = (\rho_h^m, v_h)$$

where $\zeta_h^{m+\theta} = (1-\theta)\zeta^m + \theta\zeta^{m+1}$.

We now apply Proposition 4.3 and obtain for $l = 0, \ldots, M$

$$E_l := \left\|\zeta^l\right\|^2 + C_1 k \sum_{m=0}^{l-1} \left\|\zeta_h^{m+\theta}\right\|_{\tilde{a}}^2 \leq C_2 k \sum_{m=0}^{l-1} \|\rho_h^m\|_{\tilde{*}}^2 \leq C\gamma q^{2n_G} k \sum_{m=0}^{l-1} \left\|g^{m+\theta} - \tilde{a}(\check{u}_h^m, \cdot)\right\|_{\tilde{*}}^2$$
$$\leq C'\gamma q^{2n_G} k \sum_{m=0}^{l-1} \left(\left\|g^{m+\theta}\right\|_{\tilde{*}}^2 + \|\zeta_h^m\|_{\tilde{a}}^2 + \|\tilde{u}_h^m\|_{\tilde{a}}^2\right).$$

We denote the right hand side of (4.8) with Q.

Let us first assume that $\theta = 0$ or $\theta = 1$. In this case we choose n_G large enough so that $C'\gamma q^{2n_G} \leq C_1/2$ and obtain with $l = M$

$$\left\|\zeta^M\right\|^2 + \tfrac{1}{2} C_1 k \sum_{m=0}^{M-1} \left\|\zeta_h^{m+\theta}\right\|_{\tilde{a}}^2 \leq C\gamma q^{2n_G} Q \tag{6.8}$$

since the terms $\|\zeta_h^m\|_{\tilde{a}}^2$ occur in E_M and the terms $\|\tilde{u}_h^m\|_{\tilde{a}}^2$ occur in the left hand side of (4.8).

In the general case $\theta \in [0,1]$ we use

$$\|\tilde{u}_h^m\|_{\tilde{a}}^2 \leq Ch^{-\rho}\,\|\tilde{u}_h^m\|^2 \leq Ch^{-\rho} Q$$
$$\|\zeta_h^m\|_{\tilde{a}}^2 \leq Ch^{-\rho}\,\|\zeta_h^m\|^2 \leq Ch^{-\rho} E_m$$

yielding

$$E_l \le C\gamma q^{2n_G}\Big((1+h^{-\rho})Q + k\sum_{m=0}^{l-1} h^{-\rho}E_m\Big).$$

Therefore we have estimates of the form

$$E_0 = 0, \qquad E_l \le \mu + \nu \sum_{m=0}^{l-1} E_m$$

from which we easily get by induction

$$E_l \le \mu(1+\nu)^{l-1}$$

Here we have $\nu = C\gamma q^{2n_G}h^{-\rho}T/M$. We choose n_G large enough so that $C\gamma q^{2n_G}h^{-\rho} \le 1$ and get $(1+\nu)^M \le e^T$ and

$$E_M \le C\gamma q^{2n_G}(1+h^{-\rho})Qe^T. \tag{6.9}$$

Finally we have to choose n_G large enough so that the right hand side in (6.8) or (6.9) is less than the bound in Theorem 5.4: If $k \le 1$ we have $\gamma \le Ch^{-\rho}$, and therefore we need n_G such that

$$q^{2n_G}h^{-2\rho} \le Ch^{2(p+1-\rho/2)}$$

which is satisfied for $n_G \ge C\,|\log h|$ with $C > 0$ sufficiently large, but independent of h, k. □

Theorem 6.3 allows to estimate the complexity of the time-stepping scheme with incomplete GMRES solution of the linear systems.

Corollary 6.4. *Given the compressed stiffness matrix* $\tilde{\mathbf{A}}$, *the additional work for computing* $\check{u}_h^1, \ldots, \check{u}_h^M$ *is bounded by* $CMN(\log N)^{r+1}$ *where* $r = 1$ *for* $\rho \in (0,2]$, $r = 2$ *for* $\rho = 0$.

The total work of the algorithm (for computing the compressed stiffness matrix and performing M *time steps) is bounded by* $CMN(\log N)^{r+1}$ *operations if we use exact antiderivatives, and by* $CN(\log N)^{r+d} + CMN(\log N)^{r+1}$ *operations if we use quadrature for* $d = 1, 2$.

We remark in closing that the powers of $\log N$ in these complexity estimates could be reduced by more elaborate compression techniques, see e.g. [13].

References

[1] Adams: Sobolev Spaces, Academic Press, New York 1978.

[2] L. Breiman: Probability, SIAM Publ. Philadelphia 1998.

[3] A. Cohen: Wavelet methods for operator equations. To appear in "Handbook of Numerical Analysis", 2001.

[4] W. Dahmen: Wavelet solution of operator equations. Acta Numerica, (1998).

[5] W. Dahmen and R. Stevenson: Element by element construction of wavelets satisfying stability and moment conditions, SIAM J. Numer. Anal. **37** (2000), 319-352.

[6] S. C. Eisenstat, H. C. Elman, M. H. Schultz: Variational iterative metohds for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20** (2), (1983), 345–357.

[7] L. Greengard and V. Rokhlin: A new version of the fast multipole method in 3 dimensions. Acta Numerica **6** (1997), 229-269.

[8] J.L. Lions and E. Magenes: Non-homogeneous boundary value problems and applications, Springer Verlag, Berlin (1972).

[9] T. von Petersdorff and C. Schwab: Wavelet Approximation for first kind boundary integral equations in polygons. *Numer. Math.,* **74** (1996), 479-519.

[10] T. von Petersdorff and C. Schwab: "Fully discrete multiscale Galerkin BEM", (with T. von Petersdorff) in: "Multiresolution Analysis and Partial Differential Equations", W. Dahmen, P. Kurdila and P. Oswald (Eds.), "Wavelet Analysis and its applications" Vol. 6 Academic Press, New-York (1997), 287-346.

[11] T. von Petersdorff, R. Schneider and C. Schwab: Multiwavelets for second kind integral equations, SIAM J. Num. Anal. **34** (1997), 2212-2227.

[12] P. Protter: Stochastic Integration and Differential Equations, Springer Verlag 1990.

[13] R. Schneider: Multiskalen- und Wavelet Matrixkompression. Teubner Series in Applied Mathematics, Stuttgart, 1998.

[14] L. Schwartz: Theorie des distributions, Hermann & Cie., Paris (1957).

[15] R. Stevenson: Stable three-point wavelet bases on general meshes, Numer. Math. **80** (1998), 131-158.

[16] V. Thomée: Galerkin Finite Element Methods for Parabolic Problems, Springer, Berlin Heidelberg New York, 1997