Numerical integration of damped Maxwell equations

Botchev, M.A.; Verwer, J.G.

DOI
10.1137/08072108X

Publication date
2009

Document Version
Submitted manuscript

Published in
SIAM Journal on Scientific Computing

Numerical Integration of Damped Maxwell Equations

M.A. Botchev
University of Twente, Dept. Applied Mathematics, Faculty EEMCS
P.O. Box 217, Enschede, The Netherlands
m.a.botchev@math.utwente.nl

J.G. Verwer
Center for Mathematics and Computer Science
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Jan.Verwer@cwi.nl

August 28, 2008

Abstract

We study the numerical time integration of Maxwell's equations from electromagnetism. Following the method of lines approach we start from a general semi-discrete Maxwell system for which a number of time-integration methods are considered. These methods have in common an explicit treatment of the curl terms. Central in our investigation is the question how to efficiently raise the temporal convergence order beyond the standard order of two, in particular in the presence of an explicitly or implicitly treated damping term which models conduction.

2000 Mathematics Subject Classification: Primary: 65L05, 65L20, 65M12, 65M20.
1998 ACM Computing Classification System: G.1.7, G.1.8.
Keywords and Phrases: Maxwell's equations, numerical time integration.

1 Introduction

The research reported here grew out of our interest in developing efficient numerical methods for the important Maxwell equations from electromagnetism. Maxwell's equations model the production of, and interrelations between, electric and magnetic fields and electric charge and current. The time-dependent equations appear in different forms, such as

$$\begin{aligned} \partial_t B &= -\nabla \times E\,, \\ \varepsilon \partial_t E &= \nabla \times (\mu^{-1}) B - \sigma E - J\,, \end{aligned} \tag{1.1}$$

where B is the magnetic induction flux and E the electric field. The electric current density J is a given source term and ε, μ and σ are (tensor) coefficients representing, respectively, dielectric permittivity, magnetic permeability and conductivity. The equations are posed in a three-dimensional spatial domain and provided with appropriate boundary conditions. If the equations are posed in domains without conductors, the damping term $-\sigma E$ is absent. If, in addition, the source J is taken zero, we have a prime example of a conservative wave equation system.

Numerical approximation methods for time-dependent partial differential equations (PDEs) like (1.1) are often derived in two stages (method of lines approach). First, the spatial operators are discretized on an appropriate grid covering the spatial domain, together with the accompanying boundary conditions. This leads to a time-continuous, semi-discrete problem in the form of an initial-value problem for a system of ordinary differential equations (ODEs). Second, a numerical

integration method for this ODE system is chosen, which turns the semi-discrete solution into the desired fully discrete solution on the chosen space-time grid.

In this paper we focus on the second numerical integration stage. For this purpose the paper starts from the general space-discretized Maxwell problem

$$\begin{pmatrix} M_u & 0 \\ 0 & M_v \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 & -K \\ K^T & -S \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} j_u \\ j_v \end{pmatrix}, \tag{1.2}$$

where $u = u(t)$ and $v = v(t)$ are the unknown vector (grid) functions approximating the values of the magnetic flux B and electric field E, respectively. The matrices K and K^T approximate the curl operator and the matrix S is associated with the dissipative conduction term. Throughout S can be assumed symmetric positive semi-definite. M_u and M_v are symmetric positive definite mass matrices possibly arising from a finite element or compact finite difference approximation. In case a straightforward finite difference space discretization is employed, they are diagonal or block-diagonal and thus easily inverted. The functions $j_u(t)$ and $j_v(t)$ are source terms. Typically, j_v represents the given source current J, but j_u and j_v may also contain boundary data. We do allow u and v to have different dimensions which can occur with certain finite-element methods, see e.g. [23]. Therefore, K need not to be a square matrix. The dimensions of the ODE system (1.2) are thus supposed to be as follows:

$$\begin{aligned} &u \in \mathbb{R}^m\,,\ v \in \mathbb{R}^n\,,\ \text{with}\ n \geq m\,,\ \text{and} \\ &M_u \in \mathbb{R}^{m\times m}\,,\ M_v \in \mathbb{R}^{n\times n}\,,\ K \in \mathbb{R}^{m\times n}\,,\ S \in \mathbb{R}^{n\times n}\,. \end{aligned} \tag{1.3}$$

We emphasize that the ODE system (1.2) is generic in the sense that spatial discretization of other formulations of the Maxwell equations also lead to this form. Section 4 contains an example for this observation.

In three space dimensions, (1.1) forms a system of six PDEs so that the dimensions n and m of (1.2) can take up very large values, up to 10^6 say and far beyond. Hence it is of interest to search for highly efficient methods. As mentioned, in this paper we focus on time-integration. The methods we consider do have in common an explicit treatment of the curl terms, while our central question is how to efficiently raise the temporal convergence order beyond the standard order of two, in particular in the presence of an explicitly or implicitly treated conduction term. The effectiveness of high-order time integration for finite-element solutions to conduction-free Maxwell equations has been demonstrated in [22]. The approach of [22] is based on composition methods. An attempt to extend composition ideas to problems with conductivity has been made in [25]. The proposed method is, however, restricted to scalar constant conductivity and permittivity.

The contents of the paper is as follows. Section 2 presents a stability analysis of the semi-discrete system (1.2). In particular we derive a two-by-two test model for which the numerical stability of integration methods can be examined for a wide subclass of (1.2). In Section 3 we discuss and analyze an existing, second-order integration method, which we consider as a reference method to which new, higher-order methods can be compared to assess their efficiency. This method is applied in Section 4 to a 3D Maxwell problem spatially discretized with a finite-element method. In this section we also sketch the finite-element method in some detail to illustrate the generic nature of the semi-discrete system (1.2). In Section 5 we discuss various possibilities for higher-order time integration, including explicit Runge-Kutta and composition methods and Richardson extrapolation of our second-order reference method. The Maxwell system (1.1) is a prime example of a damped wave equation system. In Section 6 we will briefly discuss another example, viz. the coupled sound and heat flow problem. For this problem the question of how to develop higher-order integration methods is closely related. Like for (1.1), we will illustrate the good performance of a second-order, symmetric composition method extrapolated to order four. Section 7 concludes the paper with final remarks.

2 Stability analysis

We begin with stability properties of the semi-discrete system (1.2). In particular we will derive a specific test model by means of which stability of integration methods can be assessed. Let $w \in \mathbb{R}^{n+m}$ denote the solution vector of (1.2) composed by u and v. Then a natural norm for (1.2) is the inner-product norm

$$\|w\|^2 = \|u\|^2_{M_u} + \|v\|^2_{M_v}\,, \qquad \|u\|^2_{M_u} = \langle M_u u, u\rangle\,, \quad \|v\|^2_{M_v} = \langle M_v v, v\rangle\,, \tag{2.1}$$

where $\langle\cdot\,,\cdot\rangle$ denotes the L_2 inner product. As S is symmetric semi-positive definite, for this norm follows

$$\frac{d}{dt}\|w\|^2 = -2\langle Sv, v\rangle \;\leq 0 \tag{2.2}$$

for the homogeneous part of (1.2), showing stability in the L_2 sense and (energy) conservation would S be zero.

For what follows it is convenient to transform (1.2) to an equivalent explicit form. For this purpose we introduce the Cholesky factorizations $L_{M_u}L^T_{M_u} = M_u$ and $L_{M_v}L^T_{M_v} = M_v$ [9] of the mass matrices M_u and M_v. The new variables $\tilde{u} = L^T_{M_u}u$ and $\tilde{v} = L^T_{M_v}v$ then satisfy the equivalent system

$$\begin{pmatrix} \tilde{u}' \\ \tilde{v}' \end{pmatrix} = \begin{pmatrix} 0 & -\tilde{K} \\ \tilde{K}^T & -\tilde{S} \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} + \begin{pmatrix} \tilde{j}_u \\ \tilde{j}_v \end{pmatrix}\,, \tag{2.3}$$

where

$$\begin{aligned} &\tilde{K} = L^{-1}_{M_u} K L^{-T}_{M_v}, \qquad \tilde{S} = L^{-1}_{M_v} S L^{-T}_{M_v}, \\ &\tilde{j}_u = L^{-1}_{M_u} j_u\,, \qquad \tilde{j}_v = L^{-1}_{M_v} j_v\,. \end{aligned} \tag{2.4}$$

Introduce the inner-product norm

$$\|\tilde{w}\|^2_2 = \|\tilde{u}\|^2_2 + \|\tilde{v}\|^2_2\,, \qquad \|\tilde{u}\|^2_2 = \langle\tilde{u},\tilde{u}\rangle\,, \quad \|\tilde{v}\|^2_2 = \langle\tilde{v},\tilde{v}\rangle\,. \tag{2.5}$$

The solution $\tilde{w}^T = [\tilde{u}, \tilde{v}]$ of the homogeneous part of (2.3) then satisfies

$$\frac{d}{dt}\|\tilde{w}\|^2_2 = -2\langle\tilde{S}\tilde{v}, \tilde{v}\rangle \;\leq 0\,, \tag{2.6}$$

while the norm is preserved under the transformation, that is, $\|\tilde{w}\|_2 = \|w\|$ and $\langle\tilde{S}\tilde{v},\tilde{v}\rangle = \langle Sv, v\rangle$. All numerical integration methods discussed later on are invariant under the transformation. So (2.3) can be used for stability analysis. We will not use the transformed system (2.3) for actual calculations.

2.1 A stability test model

If in (1.1) the conductivity coefficient σ and the permittivity coefficient ε are constant scalars instead of space-dependent tensors (3×3 matrices), then the matrices M_v and S from (1.2) are identical up to a constant for a large class of finite-element and finite-difference discretizations. That means that the matrix $\tilde{S}$ introduced in (2.3) becomes the constant diagonal matrix

$$\tilde{S} = \alpha I, \qquad \alpha = \frac{\sigma}{\varepsilon}\,. \tag{2.7}$$

This situation enables the derivation of a two-by-two system through which time-stepping stability of numerical methods for the semi-discrete system (1.2) can be examined.

The derivation starts from a second transformation based on the singular-value decomposition [9]

$$\tilde{K} = U\Sigma V^T, \tag{2.8}$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and Σ is a diagonal $m \times n$ matrix with nonnegative diagonal entries $s_1, \ldots, s_m$ satisfying

$$s_1 \geq s_2 \geq \cdots \geq s_r > s_{r+1} = \cdots = s_m = 0\,. \tag{2.9}$$

Here $r \leqslant m$ is the (row) rank of $\tilde{K}$ and the s_i are the singular values of the matrix $\tilde{K}$. The singular values of $\tilde{K}$ are just square roots of the eigenvalues of $\tilde{K}\tilde{K}^T$.

The transformed variables and source terms

$$\bar{u}(t) = U^T\tilde{u}(t)\,, \quad \bar{v}(t) = V^T\tilde{v}(t)\,, \quad \bar{j}_u(t) = U^T\tilde{j}_u(t)\,, \quad \bar{j}_v(t) = V^T\tilde{j}_v(t)\,, \tag{2.10}$$

satisfy the equivalent ODE system

$$\begin{pmatrix} \bar{u}' \\ \bar{v}' \end{pmatrix} = \begin{pmatrix} 0 & -\Sigma \\ \Sigma^T & -\alpha I \end{pmatrix} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} + \begin{pmatrix} \bar{j}_u \\ \bar{j}_v \end{pmatrix}\,, \tag{2.11}$$

where I is the $n \times n$ identity matrix. Note that the matrix transformation induced by (2.10) is a similarity transformation, so that the matrices of systems (2.3) and (2.11) have the same eigenvalues. Further, $\|\tilde{w}\|_2^2 = \|\bar{u}\|_2^2 + \|\bar{v}\|_2^2$ due to the orthogonality of U and V. Thus, if (2.7) applies, the stability of a time integration method may be studied for the homogeneous part of (2.11), provided the method is invariant under the transformation (2.10). The invariancy holds for all numerical methods discussed in the remainder of the paper.

Since the matrix Σ is diagonal, (2.11) decouples into r two-by-two systems

$$\begin{pmatrix} \hat{u}' \\ \hat{v}' \end{pmatrix} = \begin{pmatrix} 0 & -s \\ s & -\alpha \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} + \begin{pmatrix} \hat{j}_u \\ \hat{j}_v \end{pmatrix}\,, \tag{2.12}$$

with $s = s_k > 0$, $k = 1, \ldots, r$ and $n + m - 2r$ two-by-two systems

$$\begin{pmatrix} \hat{u}' \\ \hat{v}' \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -\alpha \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} + \begin{pmatrix} \hat{j}_u \\ \hat{j}_v \end{pmatrix}\,. \tag{2.13}$$

From the viewpoint of time integration, the first elementary two-by-two system (2.12) is canonical for Maxwell equation systems of which the conductivity coefficient σ and the permittivity coefficient ε are constant scalars. Note that (2.12) is equivalent to the driven oscillator equation

$$\hat{u}'' + s^2\hat{u} + \alpha\hat{u}' = d(t)\,, \qquad d(t) = \alpha\hat{j}_u + \hat{j}_u' - s\hat{j}_v\,. \tag{2.14}$$

For stability analysis we may neglect the source terms, arriving at the two-by-two stability test model

$$\begin{pmatrix} \hat{u}' \\ \hat{v}' \end{pmatrix} = \begin{pmatrix} 0 & -s \\ s & -\alpha \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}\,, \qquad s \geq 0\,, \quad \alpha \geq 0\,. \tag{2.15}$$

Stability for this test model is equivalent to stability for (2.11), which in turn is equivalent to stability for the original semi-discrete Maxwell system (1.2), provided the conductivity coefficient σ and the permittivity coefficient ε are constant scalars.

The eigenvalues of (2.15) are $(-\alpha \pm \sqrt{\alpha^2 - 4s^2})/2$. Assuming sufficiently small and large singular values s_k in (2.9), the spectra of (2.3) and (2.11) thus are cross-shaped with real eigenvalues between $-\alpha$ and 0 and complex eigenvalues with real part $-\alpha/2$ and imaginary parts $\pm\sqrt{4s_k^2 - \alpha^2}/2$.

3 A second-order reference method

From (2.2) follows that the general semi-discrete system (1.2) is either dissipative or conservative in the inner product norm introduced in (2.1). Consequently, from the viewpoint of stability, A-stable implicit Runge-Kutta methods would be ideal since these are unconditionally stable in

the inner product norm and they mimic the conservation property if, in addition, their algebraic stability matrix would be zero. This holds for the well-known Gauss methods, see the monograph [5], Sect. 4.2 or [11], Sect. IV.12 for details on this subject. However, implicit Runge-Kutta methods require the solution of linear systems which somewhat limits their practical use for large-scale systems like (1.2).

In this paper we set the possibility of using implicit Runge-Kutta methods aside and instead focus on tuned methods which only treat the damping term implicitly. For that purpose we first consider an existing method which we consider as a reference method to which new, higher-order methods should be compared to assess efficiency. The method has second order, is symmetric, treats the curl terms explicitly and the damping term implicitly. In Section 3.4 we show that the method is also free of order reduction in the presence of time-dependent boundary conditions.

3.1 The integration formula

Consider the general partitioned system

$$\begin{aligned} u' &= f(t,v)\,, \\ v' &= g(t,u,v)\,, \end{aligned} \tag{3.1}$$

and let $\tau = t_{n+1} - t_n$ denote an integration step size and u_n and v_n numerical approximations to $u(t_n)$ and $v(t_n)$. A well-known method within geometric integration, see e.g. [12, 24], is based on the composition $\Psi_\tau = \Phi_{\tau/2} \circ \Phi^*_{\tau/2}$, where Φ_τ is the (partitioned, symplectic) Euler rule

$$\begin{aligned} u_{n+1} &= u_n + \tau f(t_{n+1}, v_{n+1})\,, \\ v_{n+1} &= v_n + \tau g(t_{n+1}, u_n, v_{n+1})\,, \end{aligned} \tag{3.2}$$

and Φ^*_τ its adjoint

$$\begin{aligned} u_{n+1} &= u_n + \tau f(t_n, v_n)\,, \\ v_{n+1} &= v_n + \tau g(t_n, u_{n+1}, v_n)\,. \end{aligned} \tag{3.3}$$

Such a composition Ψ_τ thus results in the integration method

$$\begin{aligned} u_{n+1/2} &= u_n &+ \tfrac{1}{2}\tau f(t_n, v_n)\,, \\ v_{n+1} &= v_n &+ \tfrac{1}{2}\tau g(t_n, u_{n+1/2}, v_n) + \tfrac{1}{2}\tau g(t_{n+1}, u_{n+1/2}, v_{n+1})\,, \\ u_{n+1} &= u_{n+1/2} &+ \tfrac{1}{2}\tau f(t_{n+1}, v_{n+1})\,, \end{aligned} \tag{3.4}$$

which computes in a one-step manner (u_{n+1}, v_{n+1}) from (u_n, v_n). It treats f explicitly and g explicitly and implicitly with respect to its second and third argument, respectively. By construction it is symmetric and thus of second-order consistency.

This elegant composition idea applies directly to our semi-discrete Maxwell system (1.2) as this system fits in the partitioned form (3.1). Let us write

$$\begin{aligned} f(t,v) &= M_u^{-1}(-Kv + j_u(t))\,, \\ g(t,u,v) &= M_v^{-1}(K^T u - Sv + j_v(t))\,. \end{aligned} \tag{3.5}$$

Repeating the above construction then gives

$$\begin{aligned} M_u \frac{u_{n+1/2} - u_n}{\tau} &= -\tfrac{1}{2}Kv_n + \tfrac{1}{2}j_u(t_n)\,, \\ M_v \frac{v_{n+1} - v_n}{\tau} &= K^T u_{n+1/2} - \tfrac{1}{2}S(v_n + v_{n+1}) + \tfrac{1}{2}(j_v(t_n) + j_v(t_{n+1}))\,, \\ M_u \frac{u_{n+1} - u_{n+1/2}}{\tau} &= -\tfrac{1}{2}Kv_{n+1} + \tfrac{1}{2}j_u(t_{n+1})\,. \end{aligned} \tag{3.6}$$

We wish to emphasize that other choices for the t-argument are possible, in particular one which is fully compatible with the autonomous form of (3.1) (this would require two t-arguments associated

to u and v, respectively, see Section 5.2). As it is, method (3.6) suits our purpose very well. It is symmetric, has second order, an optimal error expansion (see below), and treats the curl terms explicitly and the conduction term implicitly. The mass matrices naturally give rise to implicitness such that we encounter linear system solutions for the symmetric, positive definite matrices $\frac{1}{\tau}M_u$ and $\frac{1}{\tau}M_v + \frac{1}{2}S$. Of practical importance is that the third-stage derivative computation can be copied to the first stage at the next time step so as to save computational work.

Remark 3.1 Method (3.6) is closely related to the time-staggered method

$$\begin{aligned} M_u \frac{u_{n+1/2} - u_{n-1/2}}{\tau} &= -Kv_n + j_u(t_n)\,, \\ M_v \frac{v_{n+1} - v_n}{\tau} &= K^T u_{n+1/2} - \tfrac{1}{2}S(v_n + v_{n+1}) + j_v(t_{n+1/2})\,, \end{aligned} \tag{3.7}$$

which steps from $(u_{n-1/2}, v_n)$ to $(u_{n+1/2}, v_{n+1})$, similar as in the well-known Yee-scheme [30]. Except for the source term treatment, (3.6) is in fact obtained from (3.7) by eliminating $u_{n-1/2}$ through the substitution $u_n = (u_{n-1/2} + u_{n+1/2})/2$ and by eliminating $u_{n+3/2}$ at the next time step through $u_{n+1} = (u_{n+3/2} + u_{n+1/2})/2$. This time-staggered combination of the leapfrog rule (for zero S) and implicit trapezoidal rule (for nonzero S) is well known and has for example been examined in [23]. Both methods (3.6) and (3.7) are also akin to the well-known Störmer-Verlet scheme from geometric integration, see [12], Sect. I.3.1. ◇

3.2 Energy inequalities

Stability of the linear semi-discrete systems (1.2) and (2.3) follows from the inequalities (2.2) and (2.6), respectively. It is illustrative to derive similar inequalities for method (3.6). Without source terms, for the transformed variables from (2.3) the method reads

$$\begin{aligned} \tilde{u}_{n+1/2} &= \tilde{u}_n - \tfrac{1}{2}\tau\tilde{K}\tilde{v}_n\,, \\ \tilde{v}_{n+1} &= \tilde{v}_n + \tau\tilde{K}^T\tilde{u}_{n+1/2} - \tfrac{1}{2}\tau\tilde{S}(\tilde{v}_n + \tilde{v}_{n+1})\,, \\ \tilde{u}_{n+1} &= \tilde{u}_{n+1/2} - \tfrac{1}{2}\tau\tilde{K}\tilde{v}_{n+1}\,. \end{aligned} \tag{3.8}$$

For our derivation we need to eliminate the intermediate value $\tilde{u}_{n+1/2}$ to get a scheme containing numerical solutions at whole time steps only. This is achieved by inserting $\tilde{u}_{n+1/2}$ from the first into the third line and half of it from the first and third line, respectively, into the second. After reordering we find

$$\begin{aligned} \tilde{u}_{n+1} &= \tilde{u}_n - \tfrac{1}{2}\tau\tilde{K}(\tilde{v}_n + \tilde{v}_{n+1})\,, \\ \tilde{v}_{n+1} &= \tilde{v}_n + \tfrac{1}{2}\tau\tilde{K}^T(\tilde{u}_n + \tilde{u}_{n+1}) - \tfrac{1}{4}\tau^2\tilde{K}^T\tilde{K}(\tilde{v}_n - \tilde{v}_{n+1}) - \tfrac{1}{2}\tau\tilde{S}(\tilde{v}_n + \tilde{v}_{n+1})\,. \end{aligned} \tag{3.9}$$

Now applying the inner products introduced in (2.5) yields

$$\begin{aligned} \|\tilde{u}_{n+1}\|^2 - \|\tilde{u}_n\|^2 &= -\tfrac{1}{2}\tau\langle\tilde{K}^T(\tilde{u}_n + \tilde{u}_{n+1}), \tilde{v}_n + \tilde{v}_{n+1}\rangle\,, \\ \|\tilde{v}_{n+1}\|^2 - \|\tilde{v}_n\|^2 &= \tfrac{1}{2}\tau\langle\tilde{K}^T(\tilde{u}_n + \tilde{u}_{n+1}), \tilde{v}_n + \tilde{v}_{n+1}\rangle - \\ &\quad \tfrac{1}{4}\tau^2\langle\tilde{K}(\tilde{v}_n - \tilde{v}_{n+1}), \tilde{K}(\tilde{v}_n + \tilde{v}_{n+1})\rangle - \\ &\quad \tfrac{1}{2}\tau\langle\tilde{S}(\tilde{v}_n + \tilde{v}_{n+1}), \tilde{v}_n + \tilde{v}_{n+1}\rangle\,, \end{aligned} \tag{3.10}$$

from which follows

$$\begin{aligned} \frac{(\|\tilde{u}_{n+1}\|^2 + \|\tilde{v}_{n+1}\|^2) - (\|\tilde{u}_n\|^2 + \|\tilde{v}_n\|^2)}{\tau} &= -2\,\langle\tilde{S}(\frac{\tilde{v}_n + \tilde{v}_{n+1}}{2}), \frac{\tilde{v}_n + \tilde{v}_{n+1}}{2}\rangle \\ &\quad - \frac{1}{4}\tau(\|\tilde{K}\tilde{v}_n\|^2 - \|\tilde{K}\tilde{v}_{n+1}\|^2)\,. \end{aligned} \tag{3.11}$$

This result can be seen to be the counterpart of (2.6). Likewise, the counterpart of (2.2) for the original semi-discrete system is found by the back transformation $\tilde{u}_n \to u_n$, $\tilde{v}_n \to v_n$ of this expression, giving

$$\frac{(\|u_{n+1}\|^2_{M_u} + \|v_{n+1}\|^2_{M_v}) - (\|u_n\|^2_{M_u} + \|v_n\|^2_{M_v})}{\tau} = -2\,\langle S(\frac{v_n + v_{n+1}}{2}), \frac{v_n + v_{n+1}}{2}\rangle \\ -\frac{1}{4}\tau(\langle M_u^{-1}Kv_n, Kv_n\rangle - \langle M_u^{-1}Kv_{n+1}, Kv_{n+1}\rangle)\,. \tag{3.12}$$

It follows that with a zero damping term and zero source terms we have (energy) conservation if and only if $\langle M_u^{-1}Kv_n, Kv_n\rangle = \langle M_u^{-1}Kv_{n+1}, Kv_{n+1}\rangle$, cf. (2.2). In general this will not hold. What is conserved, however, with a zero damping term and zero source terms, is the $\mathcal{O}(\tau^2)$-perturbed quantity

$$\|u_n\|^2_{M_u} + \|v_n\|^2_{M_v} - \frac{1}{4}\tau^2\langle M_u^{-1}Kv_n, Kv_n\rangle\,, \tag{3.13}$$

showing that the conservation behavior of method (3.6) is actually very good. Herewith it is of course tacitly assumed that step size τ is such that the method integrates in a stable way, something which cannot be concluded from this result due to the minus sign in front of the third term.

3.3 Test model stability

Next we will analyze the stability of method (3.6) for the test model (2.15). Let $z_\alpha = \tau\alpha \geq 0$ and $z_s = \tau s \geq 0$. Applied to this model, (3.6) yields the recurrence

$$\begin{aligned} \hat{u}_{n+1} &= \hat{u}_n - \tfrac{1}{2}z_s(\hat{v}_n + \hat{v}_{n+1})\,, \\ \hat{v}_{n+1} &= (1 - \tfrac{1}{2}z_\alpha - \tfrac{1}{2}z_s^2)\hat{v}_n + z_s\hat{u}_n - \tfrac{1}{2}z_\alpha\hat{v}_{n+1}\,, \end{aligned} \tag{3.14}$$

which we write as

$$\begin{pmatrix} \hat{u}_{n+1} \\ \hat{v}_{n+1} \end{pmatrix} = \frac{1}{1 + \frac{1}{2}z_\alpha}\begin{pmatrix} 1 + \frac{1}{2}z_\alpha - \frac{1}{2}z_s^2 & -z_s + \frac{1}{4}z_s^3 \\ z_s & 1 - \frac{1}{2}z_\alpha - \frac{1}{2}z_s^2 \end{pmatrix}\begin{pmatrix} \hat{u}_n \\ \hat{v}_n \end{pmatrix}. \tag{3.15}$$

Following common practice we call this recurrence stable if any sequence $\{(\hat{u}_n, \hat{v}_n),\, n \geq 0\}$ is bounded, which is equivalent to imposing the root condition (all roots on the unit disc and inside the disc if not simple) on the characteristic equation of the recurrence matrix.

The characteristic equation reads

$$\lambda^2 + \frac{z_s^2 - 2}{1 + \frac{1}{2}z_\alpha}\,\lambda + \frac{1 - \frac{1}{2}z_\alpha}{1 + \frac{1}{2}z_\alpha} = 0\,, \tag{3.16}$$

from which follows that for $z_\alpha = 0$ the root condition is satisfied if and only if $z_s < 2$, while for $z_\alpha > 0$ the root condition is satisfied if and only if $z_s \leq 2$. Hence we have unconditional stability for the implicitly treated conduction term and, of course, conditional stability for the explicitly treated wave terms. For $z_\alpha = 0, z_s \leq 2$ the eigenvalues have modulus one in line with the conservation property.

Corollary 3.2 Let in (1.1) the conductivity σ and permittivity ε be constant scalars and let ψ denote an eigenvalue of the matrix $\tilde{K}^T\tilde{K}$. Then method (3.6) applied to the semi-discrete Maxwell system (1.2) is stable, i.e. for $j_u = 0$, $j_v = 0$ any sequence $\{(u_n, v_n),\, n \geq 0\}$ is bounded, if and only if

$$\tau \leq \frac{2}{\sqrt{\max\psi}}\,, \tag{3.17}$$

with strict inequality for zero conduction. Note that the eigenvalues ψ of $\tilde{K}^T\tilde{K}$ coincide with those of $M_v^{-1}K^TM_u^{-1}K$ due to the similarity transformation

$$M_v^{-1}K^TM_u^{-1}K = (L_{M_v})^{-T}(\tilde{K}^T\tilde{K})(L_{M_v})^T\,. \tag{3.18}$$

Remark 3.3 When applied to test model (2.15) the time-staggered method (3.7) yields exactly the same stability restrictions. Hence Corollary 3.2 also applies to the time-staggered method. Noteworthy is that Theorem 1 in [23] is about stability of this method, but assuming a zero conduction term. That theorem states that the eigenvalues of the amplification matrix of (3.7) have unit magnitude if and only if (3.17) holds. Hence our stability result is akin to that of [23] but more general because we did not assume $S = 0$. $\diamond$

3.4 Error analysis and asymptotic expansions

Due to the symmetry the integration method (3.6) has an even global error expansion in τ. This is attractive for obtaining higher-order results through Richardson extrapolation, something we will discuss in Section 5. However, if error terms in the expansion would contain powers of K or K^T multiplying derivatives of u or v, upon space-grid refinement order reduction may occur in the case of time-dependent boundary conditions.[1)] It is therefore of interest to study and inspect the global error expansion of (3.6). We will do this for the slightly more general system

$$\begin{aligned} u' &= Ev + s_u(t)\,, \\ v' &= Bu - Sv + s_v(t)\,, \end{aligned} \tag{3.19}$$

which covers the original, semi-discrete Maxwell system (1.2) and for which we will prove that error terms only contain derivatives of u or v so that order reduction cannot occur.

Applied to system (3.19), method (3.6) becomes

$$\begin{aligned} u_{n+1/2} &= u_n + \tfrac{1}{2}\tau E v_n + \tfrac{1}{2}\tau s_u(t_n)\,, \\ v_{n+1} &= v_n + \tau B u_{n+1/2} - \tfrac{1}{2}\tau S(v_n + v_{n+1}) + \tfrac{1}{2}\tau(s_v(t_n) + s_v(t_{n+1}))\,, \\ u_{n+1} &= u_{n+1/2} + \tfrac{1}{2}\tau E v_{n+1} + \tfrac{1}{2}\tau s_u(t_{n+1})\,. \end{aligned} \tag{3.20}$$

For this system we first introduce residual local truncation errors denoted by $\delta_1, \delta_2, \delta_3$ which result from substituting true solution values. We thus write

$$\begin{aligned} u(t_{n+1/2}) &= u(t_n) + \tfrac{1}{2}\tau E v(t_n) + \tfrac{1}{2}\tau s_u(t_n) + \tau\delta_1\,, \\ v(t_{n+1}) &= v(t_n) + \tau B u(t_{n+1/2}) - \tfrac{1}{2}\tau S(v(t_n) + v(t_{n+1})) + \\ &\quad \tfrac{1}{2}\tau(s_v(t_n) + s_v(t_{n+1})) + \tau\delta_2\,, \\ u(t_{n+1}) &= u(t_{n+1/2}) + \tfrac{1}{2}\tau E v(t_{n+1}) + \tfrac{1}{2}\tau s_u(t_{n+1}) + \tau\delta_3\,, \end{aligned} \tag{3.21}$$

and Taylor expand, at $t_{n+1/2}$ for symmetry reasons, obtaining

$$\begin{aligned} \delta_1 &= \textstyle\sum_{j=2}\left(\frac{1}{(j-1)!} - \frac{1}{j!}\right)\frac{(-1)^j}{2^j}\,\tau^{j-1}u^{(j)}\,, \quad \delta_2 = \delta_4 + B\delta_5\,, \\ \delta_3 &= \textstyle\sum_{j=2}\left(\frac{1}{j!} - \frac{1}{(j-1)!}\right)\frac{1}{2^j}\,\tau^{j-1}u^{(j)}\,, \\ \delta_4 &= \textstyle\sum_{j=2'}\frac{-j}{2^j(j+1)!}\,\tau^j v^{(j+1)}\,, \quad \delta_5 = \sum_{j=2'}\frac{1}{2^j j!}\,\tau^j u^{(j)}\,, \end{aligned} \tag{3.22}$$

where $j = 2'$ means even values for j only.

Second, we introduce $\epsilon_n^u = u(t_n) - u_n$ and $\epsilon_n^v = v(t_n) - v_n$, that is, the global errors at whole time steps. Likewise we introduce the intermediate global error $\epsilon_{n+1/2}^u = u(t_{n+1/2}) - u_{n+1/2}$. Subtracting (3.20) from (3.21) then gives

$$\begin{aligned} \epsilon_{n+1/2}^u &= \epsilon_n^u + \tfrac{1}{2}\tau E\epsilon_n^v + \tau\delta_1\,, \\ \epsilon_{n+1}^v &= \epsilon_n^v + \tau B\epsilon_{n+1/2}^u - \tfrac{1}{2}\tau S(\epsilon_n^v + \epsilon_{n+1}^v) + \tau\delta_2\,, \\ \epsilon_{n+1}^u &= \epsilon_{n+1/2}^u + \tfrac{1}{2}\tau E\epsilon_{n+1}^v + \tau\delta_3\,. \end{aligned} \tag{3.23}$$

[1)] Order reduction typically occurs for one-step methods the local errors of which contain elementary differentials that not combine into higher solution derivatives. There exist quite a number of papers on order reduction by which the phenomenon is now well understood. Readers not familiar with order reduction are referred to [11], Sect. II.2.1 (where it is explained for standard Runge-Kutta methods) and references therein.

Third, we eliminate $\epsilon^u_{n+1/2}$ from the second and third line to get an error recursion which only involves errors at whole time steps. The elimination should respect the symmetry and thus result in new residuals, denoted by δ^u and δ^v, which have an even expansion in τ starting with τ^2. This is achieved by inserting $\epsilon^u_{n+1/2}$ from the first line into the third, and $\frac{1}{2}\epsilon^u_{n+1/2}$ from the first and third line, respectively, into the second. The aimed result reads, after reordering the equations,

$$\begin{aligned} \epsilon^u_{n+1} &= \epsilon^u_n + \tfrac{1}{2}\tau E(\epsilon^v_n + \epsilon^v_{n+1}) + \tau\delta^u\,, \\ \epsilon^v_{n+1} &= \epsilon^v_n + \tfrac{1}{2}\tau B(\epsilon^u_n + \epsilon^u_{n+1}) + \tfrac{1}{4}\tau^2 BE(\epsilon^v_n - \epsilon^v_{n+1}) - \tfrac{1}{2}\tau S(\epsilon^v_n + \epsilon^v_{n+1}) + \tau\delta^v\,, \end{aligned} \tag{3.24}$$

where

$$\delta^u = \delta_1 + \delta_3\,, \qquad \delta^v = \delta_2 + \frac{1}{2}\tau B(\delta_1 - \delta_3) = \delta_4 + B(\frac{1}{2}\tau(\delta_1 - \delta_3) + \delta_5)\,. \tag{3.25}$$

Inspection of δ^u and δ^v will reveal that they do possess an even expansion in τ, starting with τ^2.

Next, using a classical result on global error expansions, see e.g. [10], Sect. II.8, we let $\tau \to 0$ to recover the limit ordinary differential equation system for the global error,

$$\begin{aligned} \frac{d}{dt}\epsilon^u &= E\epsilon^v + \delta^u(t)\,, \\ \frac{d}{dt}\epsilon^v &= B\epsilon^u - S\epsilon^v + \delta^v(t)\,. \end{aligned} \tag{3.26}$$

Hereby a zero error at the initial time is assumed. At any fixed time t, we thus do have an even global error expansion for ϵ^u and ϵ^v as this holds for δ^u and δ^v. However, because B is present in δ^v, one more step is needed for cases where B is a finite difference or finite element matrix containing a negative power of a spatial grid size.

For that purpose we introduce the perturbed error $\tilde{\epsilon}^u = \epsilon^u + \frac{1}{2}\tau(\delta_1 - \delta_3) + \delta_5$. Obviously,

$$\begin{aligned} \frac{d}{dt}\tilde{\epsilon}^u &= E\epsilon^v + \tilde{\delta}^u(t)\,, \qquad \tilde{\delta}^u(t) = \delta^u(t) + \tfrac{1}{2}\tau(\delta_1'(t) - \delta_3'(t)) + \delta_5'(t)\,, \\ \frac{d}{dt}\epsilon^v &= B\tilde{\epsilon}^u - S\epsilon^v + \tilde{\delta}^v(t)\,, \quad \tilde{\delta}^v(t) = \delta_4(t)\,, \end{aligned} \tag{3.27}$$

and it follows that B has been eliminated in the new residuals $\tilde{\delta}^u$ and $\tilde{\delta}^v$, which only contain higher solution derivatives and also possess an even τ-expansion starting with τ^2. Consequently, for proper convergence behavior of $\tilde{\epsilon}^u$ and ϵ^v we only need to impose the common smoothness condition of having modestly sized solution derivatives. This then also holds for $\epsilon^u = \tilde{\epsilon}^u - \frac{1}{2}\tau(\delta_1 - \delta_3) - \delta_5$, so that grid-dependent order reduction coming from any of the matrices B, E or S cannot take place, nor for the second-order scheme, neither when extrapolating the second-order scheme to higher order in a global manner. We will discuss global (and local) Richardson extrapolation in Section 5.

Remark 3.4 The above result does not tell us anything about how the global error behaves in time, whether it growths or remains bounded. This temporal behavior is relevant in connection to long-time integration and global Richardson extrapolation. Consider a linear ODE system $\epsilon' = A\epsilon + \delta(t)$. Suppose for a certain norm $\|\cdot\|$ the stability inequality $\|e^{tA}\| \le Ce^{t\omega}$ for all $t \ge 0$, with constants $C > 0$, $\omega \in \mathbb{R}$. Then

$$\|\varepsilon(t)\| \;\le\; Ce^{t\omega}\|\varepsilon(0)\| \;+\; \frac{C}{\omega}(e^{t\omega} - 1)\max_{0\le s\le t}\|\delta(s)\|\,, \tag{3.28}$$

with convention $(e^{t\omega} - 1)/\omega = t$ in case $\omega = 0$. This well-known inequality shows that $\|\varepsilon(t)\|$ can be bounded in terms of $\|\varepsilon(0)\|$ and $\|\delta(s)\|$, $0 \le s \le t$, see e.g. [14], Sect. I.2.3.

For the Maxwell system (2.3) two cases exist. First, a zero conduction term, in which case $\omega = 0$, cf. (2.2), and linear global error growth in time is expected. Second, a non-zero conduction term, in which case $\omega < 0$ and global error built-up will be bounded. For the norm used in (2.3) the smallest possible negative ω is the negative of the smallest eigenvalue of S. Section 5.5 presents a numerical illustration of global Richardson extrapolation in a long-time integration setting. $\diamond$

4 A 3D numerical illustration

In this section we briefly sketch a spatial discretization of the 3D Maxwell equations so as to illustrate the generic form of the semi-discrete system (1.2) which for certain coefficient choices then is integrated in time by the second-order scheme (3.6). The spatial discretization is based on vector Nédélec finite elements [18, 19, 17] and is derived for the Maxwell equations with magnetic and electric fields H and E as primary variables. The formulation

$$\begin{aligned} \mu\partial_t H &= -\nabla\times E\,, \\ \varepsilon\partial_t E &= \nabla\times H - \sigma E - J\,, \end{aligned} \tag{4.1}$$

based on H and E is slightly different from (1.1) but leads to a space-discretized problem of exactly the same form. As independent variables we choose $(x,y,z)\in\Omega\subset\mathbb{R}^3$, $t\in[0,T]$ and we assume initial and boundary conditions defined by

$$E|_{t=0} = E_0(x,y,z), \qquad H|_{t=0} = H_0(x,y,z), \tag{4.2a}$$

$$(\vec{n}\times E)|_{\partial\Omega} = E_{\mathsf{bc}}, \qquad (\vec{n}\times H)|_{\partial\Omega} = H_{\mathsf{bc}}\,. \tag{4.2b}$$

The coefficients μ, ε and σ are taken constant in time and space and $\vec{n}$ denotes the outward unit normal vector to the boundary $\partial\Omega$. The boundary functions E_{bc} and H_{bc} vary in space and time.

Let $L_2(\Omega)^3$ be the space of square-integrable functions $\Omega\to\mathbb{R}^3$ and introduce

$$\mathrm{H}(\mathrm{curl},\Omega) = \left\{F\in L_2(\Omega)^3 :\ \nabla\times F\in L_2(\Omega)^3\right\}. \tag{4.3}$$

The spatial discretization is based on the following weak Galerkin formulation: find $E(x,y,z,t)$ and $H(x,y,z,t)$ in $\mathrm{H}(\mathrm{curl},\Omega)$ such that for all test functions $h,e\in\mathrm{H}(\mathrm{curl},\Omega)$ and all $t\in[0,T]$

$$\begin{aligned} \mu\partial_t(H,h) &= -(E,\nabla\times h) - \int_{\partial\Omega}(E\times h)\,\mathrm{d}\vec{s}, \\ \varepsilon\partial_t(E,e) &= (\nabla\times H,e) - \sigma(E,e) - (J,e)\,, \end{aligned} \tag{4.4}$$

with $(u,v) = \int_\Omega u\cdot v\,\mathrm{d}\omega$ being the standard inner product in $(L_2(\Omega))^3$. This formulation is discretized on a tetrahedral unstructured mesh $\Omega_h\subset\Omega$ using first-order, first-type Nédélec edge finite-element functions ϕ_j for both fields as in [13], Chapter 6. Thus the fields H, E are searched for as expansions

$$H = \sum_{j\in\Omega_h} u_j(t)\phi_j(x,y,z), \qquad E = \sum_{j\in\Omega_h} v_j(t)\phi_j(x,y,z), \tag{4.5}$$

where the summation is done through all the edges j in the mesh Ω_h. This procedure leads to a discrete weak formulation which reads in matrix form just as the ODE system (1.2) where

$$\begin{aligned} &M_u = \mu(m_{ij})\,, \quad M_v = \varepsilon(m_{ij})\,, \quad S = \sigma(m_{ij})\,, \quad m_{ij} = (\phi_j,\phi_i)\,, \\ &K = (k_{ij})\,, \quad k_{ij} = (\phi_j,\nabla\times\phi_i)\,, \\ &(j_u)_i = -\textstyle\sum_{j\in\partial\Omega_h}\left(\int_{\partial\Omega}(\phi_j\times\phi_i)\,\mathrm{d}\vec{s}\right)v_j\,, \quad (j_v)_i = -(J,\phi_i)\,, \end{aligned} \tag{4.6}$$

with zero entries $(j_u)_i$ for internal edges i and entries $(j_u)_i$ for boundary edges i defined from the boundary conditions (4.2b). We incorporate the boundary conditions (4.2b) by splitting the degrees of freedom corresponding to the known boundary values and adding them to the right hand side functions j_u and j_v. The number of equations in the resulting system then equals the doubled number of internal edges in Ω_h.

Next we give a specific example for which scheme (3.6) is used for time integration of the resulting ODE system. Let Ω be the unit cube $[0,1]^3$, let the final time $T=10$ and choose the source current $J=J(x,y,z,t)$ such that the Maxwell system (4.1) allows a specific exact solution

$$E(x,y,z,t) = \alpha(t)E_{\mathsf{stat}}(x,y,z), \qquad H(x,y,z,t) = \beta(t)H_{\mathsf{stat}}(x,y,z), \tag{4.7}$$

mesh	number of edges	longest edge $h_{\max}$	shortest edge $h_{\min}$	time step used
1	105	0.828	0.375	0.2
2	660	0.661	0.142	0.1
3	4632	0.359	0.0709	0.05
4	34608	0.250	0.0063	0.025

Table 4.1: Some mesh parameters and temporal step sizes.

where the scalar functions α, β and the vector functions E_{stat}, H_{stat} satisfy $\mu\beta'(t) = -\alpha(t)$ and $H_{\mathsf{stat}} = \nabla \times E_{\mathsf{stat}}$. The source function J is then defined as

$$J(x,y,z,t) = -(\varepsilon\alpha'(t) + \sigma\alpha(t))\, E_{\mathsf{stat}}(x,y,z) \;+\; \beta(t)\nabla \times H_{\mathsf{stat}}(x,y,z)\,, \tag{4.8}$$

and to satisfy (4.7) we choose

$$\begin{gathered} E_{\mathsf{stat}}(x,y,z) = \begin{pmatrix} \sin\pi y\, \sin\pi z \\ \sin\pi x\, \sin\pi z \\ \sin\pi x\, \sin\pi y \end{pmatrix}, \quad H_{\mathsf{stat}}(x,y,z) = \begin{pmatrix} \sin\pi x(\cos\pi y - \cos\pi z) \\ \sin\pi y(\cos\pi z - \cos\pi x) \\ \sin\pi z(\cos\pi x - \cos\pi y) \end{pmatrix}, \\ \alpha(t) = \sum_{k=1}^{3} \cos\omega_k t\,, \quad \beta(t) = -\frac{1}{\mu}\sum_{k=1}^{3} \frac{\sin\omega_k t}{\omega_k}\,, \end{gathered} \tag{4.9}$$

with $\omega_1 = 1$, $\omega_2 = 1/2$, $\omega_3 = 1/3$. Further, we take $\mu = 1$, $\varepsilon = 1$ and either $\sigma = 0$ or $\sigma = 6\pi$.

Scheme (3.6) was applied on four unstructured tetrahedral meshes of increasing size, see Table 4 for some mesh information and temporal step sizes. More information on how these meshes were generated and some mesh pictures can be found in [13]. We measured errors with respect to the exact PDE solution in the spatial L_2 integral norm at the final time $T = 10$. On the four unstructured and rather coarse meshes the errors indicate first-order convergence for H and second-order for E. First-order convergence at least complies with the theoretical convergence estimates for the implemented finite element method [18, 19, 17]. Recall that the time integration scheme is of second order.[2)] Figure 4.1 plots the errors for the two chosen values of σ for the E-field. For the damping parameter $\sigma = 6\pi$ the error appears to be notably smaller.

Finally, the finite element discretization was implemented in a Fortran code which was exported to Matlab. The mass matrices were dealt with the Matlab sparse direct solver UMFPACK with a single sparse Cholesky factorization prior to the time stepping. With these relatively coarse 3D meshes a sparse direct solver is still feasible. Needless to say that when it comes to more realistic finer 3D meshes an efficient preconditioned iterative solver is required.

5 Outlook to high-order integration

We next present an outlook to high-order integration. Like the second-order method (3.6), all aimed methods treat the curl terms explicitly. As a result, all are of comparable simplicity regarding implementation. We will focus on order four and will show that when it comes to efficiency fourth-order integration readily pays off. While the principles underlying the methods allow a still higher order, the comparative efficiency gain will necessarily decrease with the order increase. From that perspective order four is a very sensible choice. Needless to say that it is also desirable that the spatial and temporal orders match. In this section we will illustrate our ideas numerically with a 1D damped wave equation of type (1.1), which, regarding temporal dynamics, is believed to be sufficiently representative.

[2)] In a sense the second order comes for free due to the symmetry and as such is not wasteful with regard to the first-order spatial convergence for H. Increasing the spatial order is however a logical step which we plan in the near future. The need for this is also apparent in Section 5 where we discuss fourth-order temporal methods. In this paper we test these methods still with a 1D equation spatially discretized with a fourth-order compact scheme.

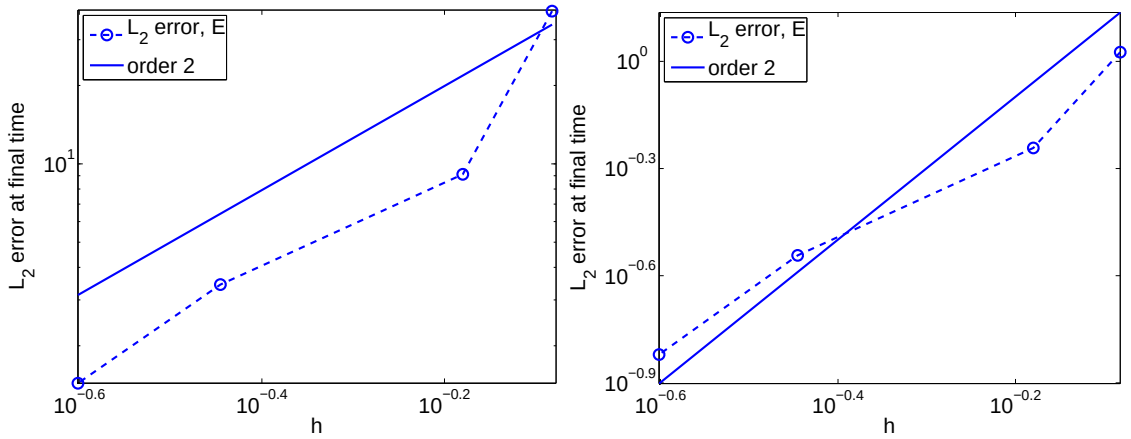


Figure 4.1: Convergence plots for the E field for $\sigma = 0$ (left) and $\sigma = 6\pi$ (right). Horizontally the mesh markers on the dashed lines correspond with the longest edge $h_{\max}$.

Any integration method which treats the curl terms explicitly does have to obey a stability restriction $\tau \sim h$ on the temporal step size, where h denotes a measure for the spatial grid size.[3)] For method (3.6) applied to test model (2.15), this is reflected by the condition $z_s = \tau s \leq 2$ from Section 3.3, where the singular values s are proportional to h^{-1}. Consequently, such methods will automatically allow considerably large conduction terms without an additional step size restriction. In this regard, the unconditional stability for the conduction term of method (3.6) and related ones can be redundant. However, having mass matrix M_u and M_v, implicit linear solution solves are necessary anyhow, in which case the unconditional stability for the conduction term comes for free.

In the following we will comment on non-stiff and stiff conduction cases. An additional motivation for this is that for related damped wave equations damping can be caused by diffusion, in which case we have of course genuine stiffness so that unconditional stability for the damping term makes sense. We return to this point in Section 6. We now discuss four different fourth-order methods, covering explicit Runge-Kutta methods, composition methods, and Richardson extrapolation.

5.1 Explicit Runge-Kutta methods

For conservative wave equations (no damping) explicit Runge-Kutta methods generate some spurious damping which might prevent one to use these methods. The conduction term alleviates this drawback for our system (1.2). Here we consider the classical, four-stage, fourth-order method (henceforth called RK4) defined by the Butcher array

$$\begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array} \tag{5.1}$$

RK4 is found in many text books and hence needs no further discussion. Of importance, however, is that RK4 requires linear system solves for (1.2) because of the mass matrices M_u and M_v. Otherwise its use is standard and well known.

[3)] In the finite difference space discretization setting, a method which overcomes this stability restriction is the ADI-FDTD (Alternating Direction Implicit – Finite Difference Time Domain) method. For practical purposes this method can be called explicit as it requires only solutions of tridiagonal linear systems. See also Remark 5.2 at the end of Section 5.3 in connection with Richardson extrapolation.

Let us consider the stability of RK4 when applied to test model (2.15) similar as we did for method (3.6) in Section 3.3. Figure 5.1 shows the (numerically determined) stability region

$$\mathcal{S} = \{(z_\alpha, z_s) : z_\alpha, z_s \geq 0 \text{ with } |\lambda| < 1, \ \lambda \text{ eigenvalues of amplification operator}\} \tag{5.2}$$

associated to test model (2.15) for which the amplification operator of RK4 is given by

$$\sum_{j=0}^{4} \frac{1}{j!} \begin{pmatrix} 0 & -z_s \\ z_s & -z_\alpha \end{pmatrix}^j . \tag{5.3}$$

Observe that $\mathcal{S}$ is symmetric around the real line and that the picture shows the upper half of $\mathcal{S}$ only (left plot). Along the vertical z_s-axis one recovers the imaginary stability interval of length $2\sqrt{2}$ of RK4 and along the horizontal z_α-axis its real stability interval of approximate length 2.78. The picture shows that $\mathcal{S}$ is sufficiently large to deal with large conduction terms, given the step size restriction $z_s \leq 2\sqrt{2}$. Of course, truly stiff terms would require no restriction on $z_\alpha > 0$, hence these cannot be dealt with.

For other higher-order explicit Runge-Kutta methods similar stability regions can be found. In line with Corollary 3.2 we thus can state the following corollary.

Corollary 5.1 Let in (1.1) the conductivity σ and permittivity ε be constant scalars and let ψ denote an eigenvalue of the matrix $M_v^{-1} K^T M_u^{-1} K$. Then any RK method applied to the semi-discrete Maxwell system (1.2) is stable, i.e. for $j_u = 0$, $j_v = 0$ any sequence $\{(u_n, v_n),\, n \geq 0\}$ is bounded, if all $(z_\alpha, z_s) \in \mathcal{S}$ where $z_\alpha = \tau\alpha$, $\alpha = \sigma/\epsilon$ and $z_s = \tau s$, $s = \sqrt{\psi}$. For example, RK4 is stable if all (z_α, z_s) are in the rectangle $0 \leq z_\alpha \leq 2.78$, $0 \leq z_s \leq 2.6$ which can be seen by inspection of the left plot of Figure 5.1. ◇

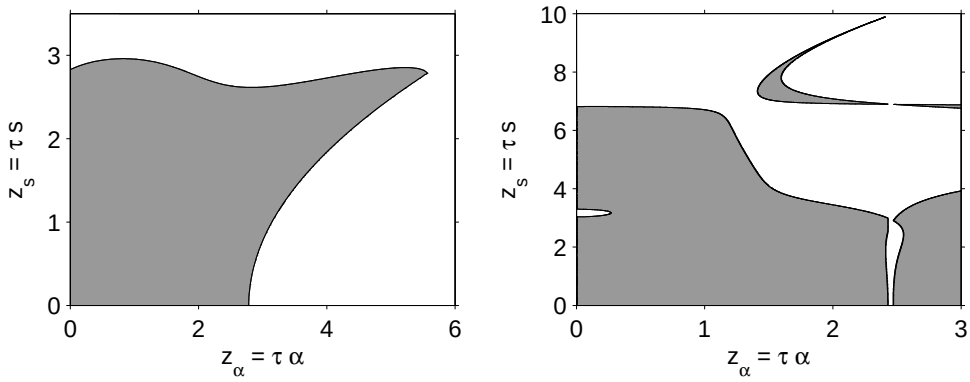


Figure 5.1: The stability region $\mathcal{S}$ for RK4 (left). Subset of $\mathcal{S}$ for CO4 (right).

5.2 Composition methods

An attractive feature of the composition technique discussed in Section 3 is that it can be extended to higher order, something which has been studied extensively within the field of geometric integration [1, 12, 15, 16, 20, 24]. Let us repeat the construction of (3.4) based on (3.1) - - (3.3), now using the symmetric composition

$$\Psi_\tau = \Phi_{\alpha_s \tau} \circ \Phi^*_{\beta_s \tau} \circ \cdots \circ \Phi_{\alpha_1 \tau} \circ \Phi^*_{\beta_1 \tau} \tag{5.4}$$

for a coefficient set α_k, β_k which is still to be chosen (see e.g. Section II.4 in [12]). Let $\alpha_0 = 0$ and define the starting step values $U_0 = u_n, V_0 = v_n$. First assume the autonomous form of (3.1). The composition Ψ_τ then can be economically and compactly written as

$$\left.\begin{aligned} U_k &= U_{k-1} + (\beta_k + \alpha_{k-1})\ \tau f(V_{k-1}) \\ V_k &= V_{k-1} + \beta_k \tau g(U_k, V_{k-1}) + \alpha_k \tau g(U_k, V_k) \end{aligned}\right\} \quad k = 1(1)s\,, \\ \begin{aligned} v_{n+1} &= V_s\,, \\ u_{n+1} &= U_s + \alpha_s \tau f(v_{n+1})\,. \end{aligned} \tag{5.5}$$

Because the final f-evaluation can be saved for the next step, s evaluations of f and g are needed per time step.

To handle non-autonomous functions $f(t, v)$ and $g(t, u, v)$ like those of (3.5) a more cumbersome notation is needed. First we introduce time levels t^v and t^u associated to v and u, respectively, and the corresponding notation $f(t^v, v)$ and $g(t^u, t^v, u, v)$. Further, we introduce new coefficients $\tilde{\alpha}_k = \alpha_1 + \cdots + \alpha_k$ and $\tilde{\beta}_k = \beta_1 + \cdots + \beta_k$ and let $\tilde{\alpha}_0 = 0$ and $\tilde{\beta}_0 = 0$. Then (5.5) becomes

$$\left.\begin{aligned} U_k &= U_{k-1} + (\beta_k + \alpha_{k-1})\ \tau f(t^v_{k-1}, V_{k-1}) \\ V_k &= V_{k-1} + \beta_k \tau g(t^u_k, t^v_{k-1}, U_k, V_{k-1}) + \alpha_k \tau g(t^u_k, t^v_k, U_k, V_k) \end{aligned}\right\} \quad k = 1(1)s\,, \\ \begin{aligned} v_{n+1} &= V_s\,, \\ u_{n+1} &= U_s + \alpha_s \tau f(t_{n+1}, v_{n+1})\,, \end{aligned} \tag{5.6}$$

where $t^v_k = t_n + (\tilde{\alpha}_k + \tilde{\beta}_k)\tau$ and $t^u_k = t_n + (\tilde{\alpha}_{k-1} + \tilde{\beta}_k)\tau$. Rewriting the Maxwell functions (3.5) as

$$\begin{aligned} f(t^v, v) &= M_u^{-1}(-Kv + j_u(t^v))\,, \\ g(t^u, t^v, u, v) &= M_v^{-1}(K^T u - Sv + j_v(t^u, t^v))\,, \end{aligned} \tag{5.7}$$

and inserting these into (5.6), then yields the counterpart of the second-order method (3.6),

$$\left.\begin{aligned} M_u \frac{U_k - U_{k-1}}{\tau} &= (\beta_k + \alpha_{k-1})(-KV_{k-1} + j_u(t^v_{k-1}))\,, \\ M_v \frac{V_k - V_{k-1}}{\tau} &= (\beta_k + \alpha_k)\, K^T U_k - S\,(\beta_k V_{k-1} + \alpha_k V_k) + \\ &\quad \beta_k j_v(t^u_k, t^v_{k-1}) + \alpha_k j_v(t^u_k, t^v_k)\,, \end{aligned}\right\} \quad k = 1(1)s\,, \\ \begin{aligned} v_{n+1} &= V_s\,, \\ M_u \frac{u_{n+1} - U_s}{\tau} &= \alpha_s \tau(-Kv_{n+1} + j_u(t_{n+1}))\,. \end{aligned} \tag{5.8}$$

The source function $j_v(t^u, t^v)$ might need a further (problem dependent) splitting into terms emanating from a physical source and terms possibly emanating from the curl discretization near the domain boundary. Generally, all time dependent terms should be temporally synchronized with all corresponding U_k and V_k values. If not, the high-order coefficient set α_k, β_k developed for the autonomous case might not give the expected order.

Aiming at order four, we have chosen $s = 5$ and

$$\begin{aligned} &\beta_1 = \alpha_5 = \tfrac{14-\sqrt{19}}{108}\,, \quad \alpha_1 = \beta_5 = \tfrac{146+5\sqrt{19}}{540}\,, \\ &\beta_2 = \alpha_4 = \tfrac{-23-20\sqrt{19}}{270}\,, \quad \alpha_2 = \beta_4 = \tfrac{-2+10\sqrt{19}}{135}\,, \quad \beta_3 = \alpha_3 = \tfrac{1}{5}\,, \end{aligned} \tag{5.9}$$

a coefficient set due to [15] which minimizes error coefficients (borrowed by us from [12], formula (V.3.6)). We used this set earlier and successfully in [29]. In the remainder of the paper we will refer to the resulting method as CO4.

The CO4 amplification operator for test model (2.15) is given by

$$\prod_{k=5}^{1} \frac{1}{1 + \alpha_k z_\alpha} \begin{pmatrix} 1 + \alpha_k z_\alpha - \alpha_k(\alpha_k + \beta_k) z_s^2 & (\alpha_k + \beta_k)(-z_s + \alpha_k \beta_k z_s^3) \\ (\alpha_k + \beta_k) z_s & 1 - \beta_k z_\alpha - \beta_k(\alpha_k + \beta_k) z_s^2 \end{pmatrix}. \tag{5.10}$$

Figure 5.1 (right plot) shows part of the associated (numerically determined) stability region (5.2), which looks quite unusual. Both for $z_\alpha = 0$ and for $z_s = 0$ we see an intermediate hole, implying that for $z_\alpha = 0$ we have the stability condition $z_s \leq 3.0$ (and eigenvalues with modulus one in line with the conservation property), while for $z_s = 0$ we have the restriction $z_\alpha < -1/\alpha_4 \approx 2.5$, approximately. On the other hand, for $z_\alpha > -1/\alpha_4$ and away from the hole, the stability region extends to infinity. The hole along the real line is due to the negative coefficient α_4 which gives a negative step size. For problems with a dissipative term this necessarily leads to instability and cuts the region in a left and right part. Similar as for RK4, for non-stiff cases this is not essential, given the inevitable restriction $z_s \leq 3.0$. Negative coefficients cannot be avoided for composition (and splitting) methods with orders beyond two [8, 26, 28]. This property thus restricts such methods to problems with small (non-stiff) dissipative terms. Note that the fourth-order time-staggered method proposed in [7] and further analyzed in [29] also has a negative step size.

Finally we note that after a proper adjustment the Corollaries 3.2 and 5.1 also hold for the current method CO4 (and other methods like those based on Richardson extrapolation discussed below).

5.3 Richardson extrapolation

The third technique we have examined is classical Richardson extrapolation of the second-order, symmetric method (3.6), henceforth called CO2. We have applied both global and local extrapolation based on the familiar extrapolation rule for symmetric methods, see e.g. [11], Sect. IV.9,

$$T_{j,k+1} = T_{j,k} + \frac{T_{j,k} - T_{j-1,k}}{(n_j/n_{j-k})^2 - 1}\,, \qquad j = 2, 3, \ldots\,, \quad k \leq j-1\,. \tag{5.11}$$

Thus, T_{j1} stands for a local or global CO2-result computed with a step size $\tau = \tau_c/n_j$ for integers $n_1 < n_2 < \cdots$ and constant base, that is coarsest, step size τ_c. Variable τ_c is allowed but is not considered. The aimed method is T_{jj} for a certain integer $j \geq 2$ giving order $2j$. As above we here restrict ourselves to order four, i.e. to $j = 2$.

5.3.1 Global extrapolation

By global extrapolation we mean passive extrapolation, hence only for output, as opposed to local extrapolation at every time step. With local extrapolation one introduces essentially a new integration method with different stability properties. With global extrapolation the stability and symmetry properties of the base method prevail. Another reason to consider global extrapolation for CO2 is its well-behaved even global error expansion for the linear system (3.19), in the sense that the expansion coefficients only contain higher solution derivatives. This implies that order reduction due to time-dependent boundary conditions cannot occur, something which does not hold with local extrapolation as we will illustrate numerically.[4)] Aiming at order four, we have chosen the most simple extrapolation T_{22} using $n_1 = 1, n_2 = 2$. Henceforth we will refer to this method as GEX4. Per base step size τ_c, GEX4 spends only three times more computations than CO2.

5.3.2 Local extrapolation

As is well known, local interpolation leads to a new integration method which might not share the good stability properties of the base method, for example loss of unconditional stability can occur, see also [11], Sect. IV.9. This indeed happens for CO2 with regard to the damping variable z_α for the harmonic sequence $n_j = j$. For this reason we choose the fourth-order local extrapolation

[4)] The fourth-order methods RK4 and CO4 and the fourth-order local extrapolation method LEX4 introduced below do suffer from order reduction when applied to semi-discrete PDEs with time-dependent Dirichlet boundary conditions.

T_{22} using $n_1 = 1, n_2 = 3$ (odd sequence), for which the unconditional stability for the damping variable z_α is preserved.[5)]

Let $\mathcal{M}(z_\alpha, z_s)$ denote the amplification operator of CO2 occurring in (3.15). The amplification operator of T_{22} then is given by $\frac{9}{8}\mathcal{M}^3(z_\alpha/3, z_s/3) - \frac{1}{8}\mathcal{M}(z_\alpha, z_s)$. The left plot of Figure 5.2 shows the associated stability region $\mathcal{S}$ (with z_a restricted to $0 \le z_\alpha \le 30$) which extends to infinity along the z_α-axis. The right plot zooms in near the vertical axis, showing that the stability interval for z_s amounts to $0 \le z_s \le 2.85$, approximately.

Henceforth we will refer to the current local extrapolation method as LEX4. Per base step size τ_c, LEX4 spends 4.5 times more computations than CO2, rather than 4, because in one application of CO2 its third-stage computation cannot be passed on to the next time step. Hence per base step size LEX4 needs 50% more computations than GEX4. Also note that LEX4 does not preserve the symmetry of CO2, so that for $z_\alpha = 0$ the moduli of the eigenvalues of the amplification operator do deviate from one along the stability interval for z_s. For conduction-free problems this incurs spurious damping of higher harmonics. This effect is restricted, however, to only the truly higher harmonics, since the moduli do stay very close to one on a significant part of the stability interval, see Figure 5.3.

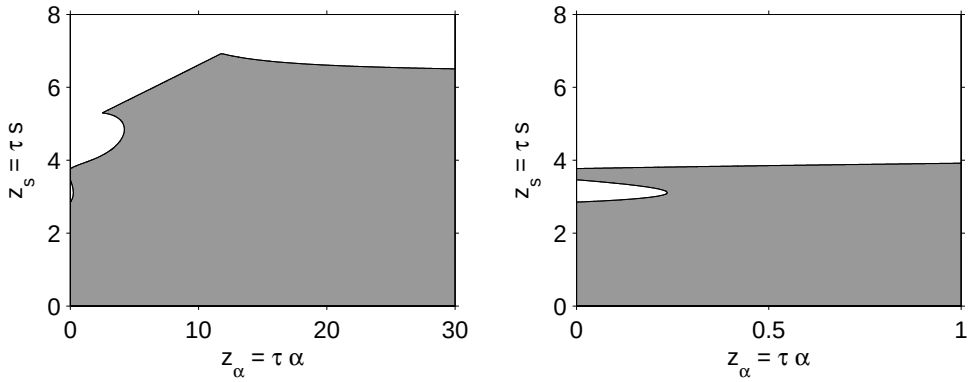


Figure 5.2: The stability region $\mathcal{S}$ of method LEX4 (left plot). The right plot zooms in near the vertical axis.

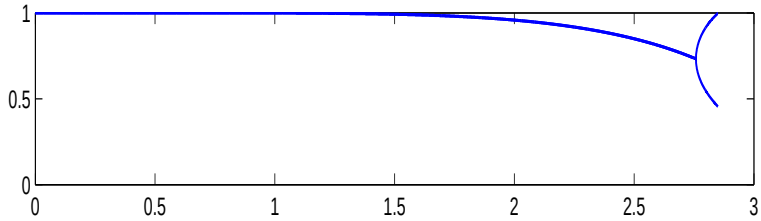


Figure 5.3: The eigenvalue moduli of method LEX4 along the stability interval $0 \le z_s \le 2.85$.

Remark 5.2 One of the referees brought [6] to our attention where for Maxwell's equation local and global Richardson extrapolation is discussed for the ADI-FDTD (Alternating Direction Implicit – Finite Difference Time Domain) method. Increasing the order of this method while maintaining its very good stability has been successfully illustrated. In addition extrapolation is discussed for wave equations in a method of lines setting for the implicit trapezoidal rule and the

[5)] For the harmonic sequence standard smoothing as e.g. discussed in [11], Sect. IV.9, will restore unconditional stability in z_α (see also [3] for a different approach). However, due to the additional costs for smoothing we expect that T_{22} for $n_1 = 1, n_2 = 2$ with smoothing will not lead to a better method than T_{22} for $n_1 = 1, n_2 = 3$ without smoothing. For higher orders smoothing may be of interest.

GBS (Gragg-Bulirsch-Stoer) scheme. Regarding extrapolation we focus on method CO2, that is method (3.6), taking into account damping terms giving dissipative effects and time-dependent boundary conditions giving order reduction effects. For the ADI-FDTD method such effects are not discussed in [6] as periodicity is assumed for Maxwell's equation without damping terms. Also note that this method is not applicable to the unstructured grid vector finite element discretization considered in Section 4. ◇

5.4 Numerical illustration

Although simple, the 1D damped linear wave equation

$$B_t = E_x\,, \quad E_t = B_x - \alpha E + \alpha\psi(x,t)\,, \qquad 0 \le x \le 1\,, \quad t > 0\,, \tag{5.12}$$

serves our purpose. We let $\psi(x,t) = E(x,t)$ so that (5.12) has the generic solution

$$B(x,t) = \frac{1}{2}(B_0(x+t) \,+\, B_0(x-t))\,, \quad E(x,t) = \frac{1}{2}(B_0(x+t) \,-\, B_0(x-t))\,, \tag{5.13}$$

where we choose B_0 as the pulse profile $B_0(x) = e^{-100(x-\frac{1}{2})^2}$. See Figure 5.4, which shows that for $t \le 0.1$ we have numerically zero boundary values, while for later times the boundary values become time dependent. As in [29] we use this to illustrate the order reduction phenomenon.

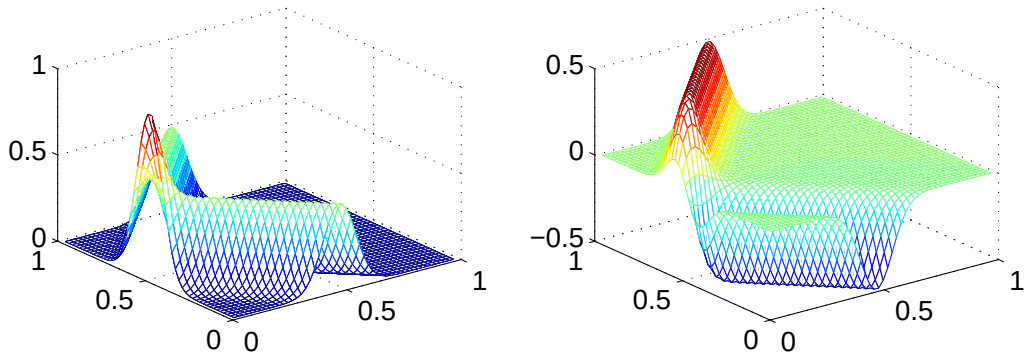


Figure 5.4: The exact solution of the 1D test problem (5.12)-(5.13). At the left B, at the right E.

Let $h = 1/(N+1)$, $x_i = ih$ for $i = 0, 1, \ldots, N+1$, and let $u_i(t)$ and $v_i(t)$ denote the semi-discrete approximations to $B(x_i,t)$ and $E(x_i,t)$, respectively. We then discretize $B_t = E_x$ in space with the fourth-order compact scheme

$$\frac{1}{6}(u'_{i-1} + 4u'_i + u'_{i+1}) = \frac{1}{2h}(v_{i+1} - v_{i-1})\,, \quad i = 1\,, \ldots, N\,. \tag{5.14}$$

The boundary values u'_0, u'_{N+1} and v_0, v_{n+1} are prescribed from the exact solution.[6)] The second equation is semi-discretized similarly. Arranging the unknowns u_i, v_i in vectors u, v of length N, we then arrive at a semi-discrete system which fits in class (1.2) and to which the stability analysis of Section 2 applies, revealing a maximal singular value $s \approx 1.74/h$.

[6)] Because the solution consists of outgoing waves and is defined for all x, imposing this boundary condition is a consequence of the finite spatial domain and the specific spatial scheme. It does serve our purpose however on illustrating the order reduction phenomenon.

We choose the damping coefficient $\alpha = 1$ (non-stiff case), so that the following critical step sizes for stability emanating from the wave terms apply,

$$\tau_c = \begin{cases} 2.0h/1.74 & \approx 1.14h & \text{CO2 and GEX4} \\ 2\sqrt{2}h/1.74 & \approx 1.62h & \text{RK4} \\ 2.85h/1.74 & \approx 1.63h & \text{LEX4} \\ 3.0h/1.74 & \approx 1.72h & \text{CO4} \end{cases} \tag{5.15}$$

With a minor adjustment to hit chosen output times within an integer number of steps, step sizes (5.15) are used in the numerical tests. Figure 5.5 contains convergence results at $t = 0.1$ (left) and $t = 0.5$ (right). The marks correspond with $N = 40, 80, \ldots, 1280$ and since $h = 1/(N+1)$ and $\tau \approx \tau_c$, τ and h decrease simultaneously. Hence we look at PDE convergence rather than ODE convergence for h fixed. The loglog plots show efficiency. That is, we plot maximum norm errors for B (PDE solution minus fully discrete solution over all components of u) versus computational work (number of time steps times number of (f, g)-evaluations per step times number of spatial grid points). For component E the errors are alike. Recall that for CO2, GEX4, RK4 and CO4 the numbers of (f, g)-evaluations per step amount to, respectively, one, three, four and five, while for LEX4 it is four and a half.

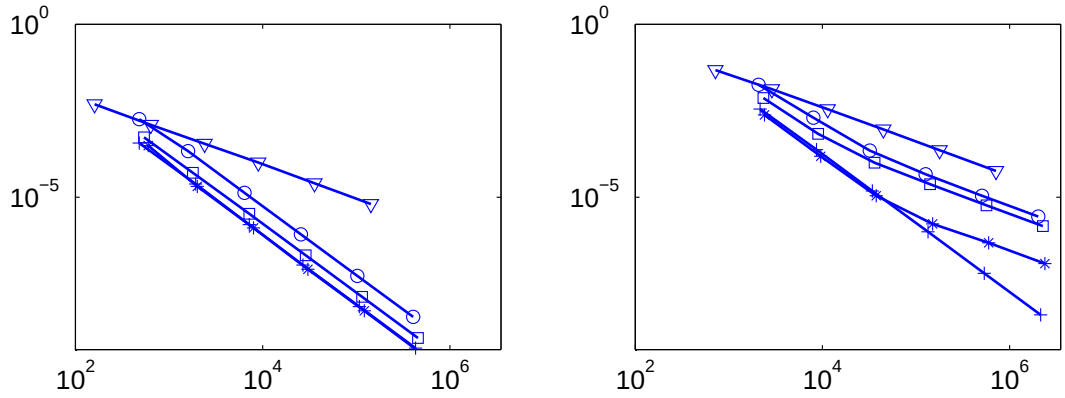


Figure 5.5: Loglog convergence (error versus work) plots for problem (5.12) at $t = 0.1$ (left) and $t = 0.5$ (right). Non-stiff case $\alpha = 1$. CO2 $\triangledown$-marks, GEX4 $+$-marks, LEX4 $\square$-marks, RK4 o-marks, CO4 $*$-marks.

Both plots clearly show the efficiency advantage of higher order integration, distinguishing between CO2, which shows its second-order convergence[7)], and the other four methods. At $t = 0.1$ all four fourth-order methods do converge for $\tau, h \to 0$ with their ODE order four. No reduction occurs because up to $t = 0.1$ the boundary values are numerically zero. On the other hand, at $t = 0.5$ we clearly observe reduction of the PDE order for RK4, LEX4 and CO4, down to order two, while GEX4 is free from reduction. The latter observation is in accordance with the error analysis presented in Section 3. The effect of reduction becomes visible on fine grids only, which is due to the fact that error terms causing reduction generally have relatively small error coefficients. This holds in particular for CO4. We emphasize that the reduction is so clearly visible because we decrease τ and h simultaneously. Would we fix h, reduce τ, and compare with an exact ODE solution, eventually the ODE order four will be found, accompanied however with comparatively large errors caused by those error terms causing reduction. Finally, among the fourth-order methods, RK4 is the least efficient method. In the absence of order reduction GEX4 and CO4 are equally efficient, indicating that the spatial error dominates. Overall GEX4 is the winner in the current test.

[7)] Because of its temporal order two, the fourth-order spatial discretization is a bit of a waste for CO2 as fourth-order spatial discretization is more expensive than the most simple second-order one. In this regard CO2 is not equally treated in this comparison. But also with second-order in space it will be less efficient than its three fourth-order competitors.

5.5 Local versus global Richardson extrapolation

It should be emphasized that in the experiment with time-dependent Dirichlet boundary values the significantly better performance of global compared to local extrapolation is due to order reduction. Without boundary effects, generally there will not be much difference in performance, at least for dissipative problems for which global error built-up is bounded, cf. Remark 3.4. However, for non-dissipative problems and long-time integration one readily observes linear global error built-up to the extent that leading error terms have become too large to be efficiently annihilated by global extrapolation. If this occurs, global extrapolation will become lesser and lesser efficient over time, simply because the extrapolation is then delayed too long.

To illustrate this, we have solved, with and without damping, the inhomogeneous test model (2.12) for the particular solution $\hat{u}(t) = \sin(2\pi t)$, $\hat{v}(t) = -\frac{2\pi}{s}\cos(2\pi t)$. For $\alpha = 0$ (no damping) and $s = 1$ we have applied method LEX4 and GEX4 over the interval $[0, 3.0 \cdot 10^4]$. Figure 5.6 shows the absolute errors in $\hat{u}$, measured after every 50-th period of 2π. For simplicity of testing, equal base step sizes $\tau_c = 1/15$ were used.

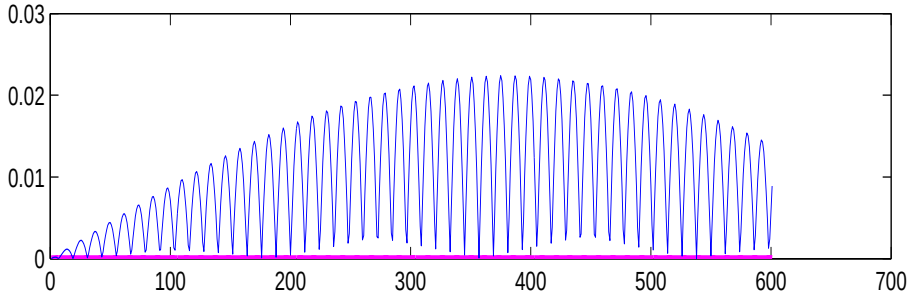


Figure 5.6: The driven oscillator test. Vertical axis: absolute errors in $\hat{u}$. Horizontal axis: $t/50$.

The errors reveal a truly different behavior. The (oscillatory) error for the local method LEX4 remains small (the thick line at the bottom), with a more or less constant maximum of about $3.1 \cdot 10^{-5}$. On the other hand, the oscillatory error for the global method GEX4 first growths significantly, with its maximum linearly up to about $2.2 \cdot 10^{-2}$, and then decreases again (the numerical solution becomes in phase with the exact solution). The linear growth for GEX4 is due to linear growth exhibited by its based method CO2 and is simply copied by the global, passive extrapolation to higher order. On the other hand, the local method LEX4 annihilates leading error terms at each time step and introduces artificial damping. In the present test this damping is minor, but it does help to counteract linear global error growth.

In our second test, which serves to illustrate the behavior with damping, we have chosen $\alpha = 1$. In this test both methods give nearly equal errors (not shown here), with maxima almost constant in time and equal to $6.3 \cdot 10^{-6}$, approximately. The damping gives rise to a bounded global error over time for the base method CO2, resulting in this very similar performance. Would we have incorporated the difference in costs per step by choosing a smaller step size τ for GEX4, in this special case the global method would even be a factor $(3/2)^4$ more accurate for equal work. For another chosen sequence n_j introduced in (5.11), we would have a different factor of course.

For a further comparison between global and local extrapolation we have done similar tests with the symmetric method (3.4) applied to Kepler's two-body problem which is often used as a test problem in geometric integration research, see e.g. [12]. For this nonlinear, conservative (Hamiltonian) problem, the difference in accuracy between global and local extrapolation is truly more significant than for the driven oscillator problem and is strongly in favor of the local approach (results are not shown here). For a thorough analysis of combining composition and local extrapolation to raise the order of geometric integrators we refer the interested reader to [2, 4].

6 The coupled sound and heat flow problem

The Maxwell equations (1.1) provide a prime example of a damped wave equation system. This suggests that the integration methods we discussed may be applicable to other damped wave equations as well. In this section we will illustrate this. As an example we use the coupled sound and heat flow problem, while focusing on an efficient combination of second-order, symmetric composition and global and local Richardson extrapolation, similar to what we did for methods CO2 and GEX4 and LEX4. We focus on extrapolation methods because the coupled sound and heat flow problem contains the Laplace operator. This gives rise to infinite stiffness, ruling out explicit Runge-Kutta methods like RK4 and composition methods like CO4. We expect no significant difference between local and global extrapolation because the problem at hand is dissipative and is formulated with periodic boundary conditions so that order reduction effects play no role.

We consider the scaled linearized equations from [21], Sect. 10.4,

$$\begin{aligned} e_t &= d\,\Delta e - c\,\nabla\cdot\mathbf{u}\,, \\ v_t &= c\,\nabla\cdot\mathbf{u}\,, \\ \mathbf{u}_t &= c\,\nabla v - c(\gamma-1)\,\nabla e\,, \end{aligned} \tag{6.1}$$

expressing, respectively, conservation of energy, mass and momentum, and wherein $v, \mathbf{u}$ and e represent specific volume, material velocity and specific internal energy; c is the isothermal sound speed, $\gamma > 1$ the ratio of specific heat, and $d \geq 0$ the thermal conductivity coefficient. This time-dependent PDE system is posed in a one-, two-, or three-dimensional space domain and should be provided with boundary conditions. For convenience of presentation we suppose periodic boundary conditions.

Of importance is that the damping term $d\Delta e$ gives rise to infinite stiffness, suggesting an implicit treatment, as opposed to the remaining wave terms which all can be treated explicitly. Introduce the notation $e_t = f(\mathbf{u}, e), v_t = g(\mathbf{u}), \mathbf{u}_t = h(v, e)$ and the Euler-type scheme

$$\Phi_\tau \left\{ \begin{array}{lcl} e_{n+1} &=& e_n + \tau f(\mathbf{u}_n, e_{n+1})\,, \\ v_{n+1} &=& v_n + \tau g(\mathbf{u}_n)\,, \\ \mathbf{u}_{n+1} &=& \mathbf{u}_n + \tau h(v_{n+1}, e_{n+1})\,, \end{array} \right. \tag{6.2}$$

formulated at the PDE level. The composition $\Psi_\tau = \Phi_{\tau/2} \circ \Phi^*_{\tau/2}$ then defines the symmetric, second-order, one-step integration method

$$\begin{array}{lcll} \mathbf{u}_{n+1/2} &=& \mathbf{u}_n & + \ \frac{1}{2}\tau\,(c\,\nabla v_n - c(\gamma-1)\,\nabla e_n)\,, \\ v_{n+1} &=& v_n & + \ \tau\,c\,\nabla\cdot\mathbf{u}_{n+1/2}\,, \\ e_{n+1} &=& e_n & + \ \frac{1}{2}\tau\,d\,(\Delta e_n + \Delta e_{n+1}) - \tau\,c\,\nabla\cdot\mathbf{u}_{n+1/2}\,, \\ \mathbf{u}_{n+1} &=& \mathbf{u}_{n+1/2} & + \ \frac{1}{2}\tau\,(c\,\nabla v_{n+1} - c(\gamma-1)\,\nabla e_{n+1})\,, \end{array} \tag{6.3}$$

which uses effectively 3 stages per step because the fourth stage can be copied to the next time step. The method is explicit in velocity and volume and implicit-explicit in energy so as to cope with the infinitely stiff damping term $d\Delta e$.

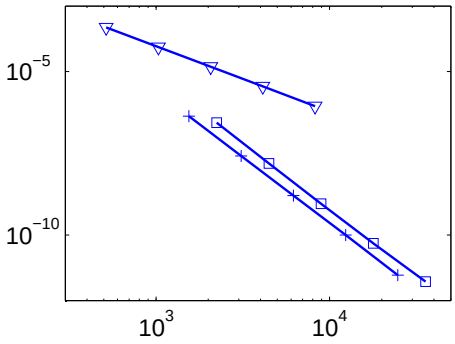
Similar as we discussed before, τ^2-extrapolation is applicable due to the symmetry. To illustrate that this can be very efficient, we have applied (6.3) as base method (the counterpart of CO2), its fourth-order global extrapolation T_{22}, using $n_1 = 1, n_2 = 2$ (the counterpart of GEX4), and its fourth-order local extrapolation T_{22}, using $n_1 = 1, n_2 = 3$ (the counterpart of LEX4). The test problem from class (6.1) is two-dimensional and is borrowed from [27]. This problem is defined on the unit square, has periodic boundary conditions and time interval $0 \leq t \leq 1$. Further, it has as initial velocities in x and y direction the periodic functions $u_1 = \frac{2}{5}\pi \sin^2(\pi x)\,\sin(2\pi y)$ and $u_2 = -\frac{1}{5}\pi \sin^2(\pi y)\sin(2\pi x)$ and has a zero initial field for v and e. The problem coefficients are given by $c = 1, \gamma = 3$ and d is the peaked function $d = \frac{1}{10}\sin^{10}(\pi x)\,\sin^{10}(\pi y)$. With this setup we encounter, after linearization, real negative eigenvalues and eigenvalues very close to

the imaginary axis since d is close to zero on part of the domain. Fourth-order central space discretization was used on a single uniform grid with grid size $h = 1/100$. The linear systems of algebraic equations arising from the Laplace operator were solved by LU-decomposition. We refer to [27] for more details, amongst others for how to carry out Fourier-von Neumann stability analysis to the space-discretized version of (6.3).

The figure below gives an efficiency-accuracy plot for the step sizes

$$\tau_c = (1\,,\,\frac{1}{2}\,,\,\ldots\,,\,\frac{1}{16}\,)\,\tau_h \quad \text{with} \quad \tau_h = \begin{cases} 2.00\cdot(5h/7c\sqrt{2\gamma}) \;\approx\; 0.0058 & \text{CO2 and GEX4} \\ 2.85\cdot(5h/7c\sqrt{2\gamma}) \;\approx\; 0.0083 & \text{LEX4} \end{cases} \tag{6.4}$$

selected on the basis of the Fourier-von Neumann stability analysis. Along the vertical axis we plot the maximum absolute error at $t = 1$ chosen with respect to a highly accurate reference ODE solution, and along the horizontal axis the computational work expressed as numbers of time steps times stages (three for the base method (6.3)). The loglog plot with $+$-marks for global and $\square$-marks for local extrapolation is self evident, showing a huge efficiency gain for extrapolation when high accuracy is wanted.



7 Final remarks

A question of general numerical interest is whether for partial differential equations high-order discretization methods are more efficient than more commonly used second-order ones. The current paper was devoted to high-order time-stepping methods for damped Maxwell equations (1.1). We have shown that if the curl terms can be treated explicitly, a variety of high-order techniques prove very useful, including explicit Runge-Kutta methods, symmetric composition methods, and Richardson extrapolation based on the second-order symmetric method (3.6). We have also analyzed stability of these methods for the full Maxwell system (1.2) in the case of constant conductivity and permittivity.

While the Runge-Kutta and high-order composition methods are restricted to 'non-stiff' damping terms, this restriction does not hold for the extrapolation approach, the success of which to a great extent is due to the bonus of τ^2-extrapolation. The Richardson extrapolation technique is classic and symmetric methods like (3.6) are well-known, yet successfully combining the two for solving damped wave equations like (1.1) and related damped wave equations like (6.1) has got little attention as far as we know. We have applied extrapolation both locally and globally. The most interesting feature of the global approach, for method (3.6), is that it does not suffer from order reduction for problems with time-dependent Dirichlet boundary conditions. Of course, necessary for temporal global extrapolation to work well is, besides a sufficiently smooth global error expansion, limited error built-up. This is the case for the large class of dissipative problems. However, when order reduction is no issue, we would in general yet advocate the local approach as then leading error terms are eliminated instantaneously. Needless to say, to justify a high temporal order also the spatial discretization order should be high enough.

Finally, interested readers should also consult [6] concerning the successful application of Richardson extrapolation to the ADI-FDTD method mentioned in Remark 5.2.

Acknowledgement M.A. Botchev acknowledges financial support from the Dutch government through the BSIK ICT project BRICKS, subproject MSV1 Scientific Computing.

References

[1] S. Blanes, P.C. Moan (2002), *Practical symplectic partitioned Runge-Kutta and Runge-Kutta-Nyström methods*. J. Comp. Appl. Math. 142, pp. 313–330.

[2] S. Blanes, F. Casas (2005), *Raising the order of geometric numerical integrators by composition and extrapolation*. Numerical Algorithms 38, pp. 305–326.

[3] R.P.K. Chan (1993), *Generalized symmetric Runge-Kutta methods*. Computing 50, pp. 31–49.

[4] R.P.K. Chan, A. Murua (2000), *Extrapolation of symplectic methods for Hamiltonian problems*. Appl. Numer. Math. 34, pp. 189–205.

[5] K. Dekker, J.G. Verwer (1984), *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*. North-Holland, Amsterdam.

[6] B. Fornberg, J. Zuev, J. Lee (2007), *Stability and accuracy of time-extrapolated ADI-FDTD methods for solving wave equations*. J. Comp. Appl. Math. 200, pp. 178–192.

[7] M. Ghrist, B. Fornberg, T.A. Driscoll (2000), *Staggered time integrations for wave equations*. SIAM. J. Numer. Anal. 38, pp. 718–741.

[8] D. Goldman, T.J. Kaper (1996), *N-th order split operator schemes and non-reversible systems*. SIAM J. Numer. Anal. 33, pp. 349–367.

[9] G.H. Golub, C.F. van Loan (1996), *Matrix Computations*. Third edition. John Hopkins Univ. Press, Baltimore.

[10] E. Hairer, S.P. Nørsett, G. Wanner (1993), *Solving Ordinary Differential Equations I – Nonstiff Problems*. Second edition, Springer Series in Computational Mathematics, Vol. 8, Springer-Verlag, Berlin.

[11] E. Hairer, G. Wanner (1996), *Solving Ordinary Differential Equations II – Stiff and Differential-Algebraic Problems*. Second edition, Springer Series in Computational Mathematics, Vol. 14, Springer-Verlag, Berlin.

[12] E. Hairer, C. Lubich, G. Wanner (2002), *Geometric Numerical Integration*. Springer Series in Computational Mathematics, Vol. 31, Springer.

[13] D. Harutyunyan (2007), *Adaptive vector finite element methods for the Maxwell equations*. PhD thesis, University of Twente. Available at http://eprints.eemcs.utwente.nl/9859/

[14] W. Hundsdorfer, J.G. Verwer (2003), *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Series in Computational Mathematics, Vol. 33, Springer, Berlin.

[15] R.I. McLachlan (1995), *On the numerical integration of ordinary differential equations by symmetric composition methods*. SIAM J. Sci. Comput. 16, pp. 151–168.

[16] R.I. McLachlan, G.R.W. Quispel (2002), *Splitting methods*. Acta Numerica 2002, pp. 341–434.

[17] P. Monk (2003), *Finite Element Methods for Maxwell's Equations*. Oxford University Press.

[18] J.-C. Nédélec (1980), *Mixed finite elements in* $\mathbf{R}^3$. Numer. Math., pp. 315–341.

[19] J.-C. Nédélec (1986), A new family of mixed finite elements in $\mathbf{R}^3$. *Numer. Math.*, 50(1):57–81.

[20] A. Murua, J.M. Sanz-Serna (1999), *Order conditions for numerical integrators obtained by composing simpler integrators*. Philos. Trans. Royal Soc. A 357, pp. 1079–1100.

[21] R.D. Richtmyer, K.W. Morton (1967), *Difference Methods for Initial-Value Problems*. Second edition, John Wiley & Sons, Interscience Publishers, New York.

[22] R. Rieben, D. White, G. Rodrigue (2004), *High-order symplectic integration methods for finite element solutions to time dependent Maxwell equations*. IEEE Transactions on Antennas and Propagation 52, pp. 2190–2195.

[23] G. Rodrigue, D. White (2001), *A vector finite element time-domain method for solving Maxwell's equations on unstructured hexahedral grids*. SIAM J. Sci. Comput. 23, pp. 683–706.

[24] J.M. Sanz-Serna, M.P. Calvo (1994), *Numerical Hamiltonian Problems*. Chapman & Hall, London.

[25] Z. Shao, Z. Shen, Q. He, G. Wei (2003), *A generalized higher order finite-difference time-domain method and its application in guided-wave problems*. IEEE Transactions on Microwave Theory and Techniques, 51, pp. 856–861.

[26] Q. Sheng (1989), *Solving partial differential equations by exponential splittings*. IMA J. Numer. Anal. 9, pp. 199–212.

[27] B.P. Sommeijer, J.G. Verwer (2007), *On stabilized integration for time-dependent PDEs*. J. Comput. Phys. 224, pp. 3–16.

[28] M. Suzuki (1990), *Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations*. Phys. Lett. A 146, pp. 319–323.

[29] J.G. Verwer (2007), *On time staggering for wave equations*. Journal of Scientific Computing 33, pp. 139–154.

[30] K.S. Yee (1966), *Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media*. IEEE Trans. Antennas Propag. 14, pp. 302–307.