An enhancement of instruments for solution of general transmission line equations with method of lines, impedance-/admittance and field transformation in combination with finite differences.

Waldemar Spiller (walspill@yahoo.de)
FernUniversitat in Hagen https://orcid.org/0000-0001-6365-4958

Research Article

Keywords:

Posted Date: July 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-901151/v2

Optical and Quantum Electronics manuscript No.
(will be inserted by the editor)

An enhancement of instruments for solution of general transmission line equations with method of lines, impedance-/admittance and field transformation in combination with finite differences

Waldemar Spiller

Received: date / Accepted: date

Abstract The Method of Lines (MoL) in combination with impedance/admittance and field transformation (IAFT) is used to analyze electromagnetic waves. The used cases are waveguiding structures in microwave technology and optics. The core of the theory is the solution of generalized transmission line equations (GTL). In the case of complex structures, a combination with finite differences (FD) can be used. The quality of this solution essentially depends on the effectiveness of the used interpolation of the differences. The individual steps of the FD are permanently linked to the steps of the fully vectorial impedance/admittance and field transformation, so standard libraries cannot be used. Two approaches based on the linear and quadratic interpolation were built into the impedance/admittance and field transformation in the past. However, the degree of improvement due to one or another kind of interpolation depends on the concrete behavior of the solution sought. In the case of complex structures, choosing the appropriate type of interpolation should be an effective aid. In this paper, an extension of the family of built-in methods is proposed - with the possibility of being able to build any known numerical method from the class of one-step or multi-step methods into the GTL solution. These can be higher-order methods, including fast explicit methods, or particularly stable implicit methods. The transmission matrices for the impedance/admittance and field transformation serve as the building site. To illustrate the procedure, some different methods are integrated into the GTL solution. The efficiency of the solutions is tested on some test structures and compared with each other and with existing solutions. The relevant waveguide specifics were discussed. Initial systematics and recommendations for users were derived.

Keywords Method of lines, generalized transmission line equations, impedance/admittance transformation, waveguide structures, finite differences, finite differences with second-order accuracy.

1 Introduction

The Method of Lines (MoL) is a semi-analytical versatile tool for the solution of partial differential equations (Helfert and Pregla, 2002). In general, one can differentiate between two numerical approaches to realizing the solutions. In the first, very common approach, the differential equations are solved directly by temporarily "freezing" the partial derivatives, e.g., (Schiesser, 1991), (Hamdi et al., 2007), (Schiesser and Griffiths, 2009). In the second approach discussed in this paper, the solution is more oriented towards the purpose of analyzing electromagnetic waves, e.g., (Helfert and Pregla, 2002),(Pregla and Helfert, 2002), (Helfert et al., 2003), (Barcz, Helfert and Pregla, 2002), (Pregla and Pascher, 1989), (Pregla, 2008) etc.

The MoL, in the sense of the second approach, is known for a long time. However, in the last 10 years, only extremely few applications or further developments have become known. This seems particularly surprising because this approach has outstanding properties, e.g. stationary behavior, relation to the transmission line theory, (Chen, 2004) ch.V, no spurious modes, stability and many others, (Pregla, 2008). One can only speculate that this excellent theory may seem less applicable to potential users. The main aim of this paper is to expand the usability of the MoL-IAFT-FD, an extension and possibly a flexibilization of the MoL instruments. The usability is considered in terms of efficiency (computational time compared to accuracy), implementation effort, portability to another structure and stability. The main idea is that for specific complex structures, by using certain methods, the computational effort can be reduced significantly, e.g., with the same accuracy. A reduction in the calculation effort is important for the MoL-IAFT-FD also because of the recursive character of the calculations: This reduces

W. Spiller, E-mail: walspill@yahoo.de

the probability of a cumulative accumulation of rounding errors $O(1/\Delta\overline{u})$ and thus a possible divergence of the solution, (Bronstein et al, 2005).

The first possibility of improvement results from the comparison of the existing instruments of the MoL-IAFT-FD with the numerous instruments of the numerics of the ordinary differential equations (ODE). The numerics of the ODE have several method classes available for the user to choose from, e.g., one-step and multistep approaches, which in turn contain numerous methods of integrating the ODE, e.g. explicit or implicit methods of different orders of accuracy. Thus, a wide range of tasks can be calculated efficiently, as the user has multiple choices. In contrast to this, the MoL-IAFT-FD has only two methods of numerical integration of the ODE: the linear and the quadratic interpolation, (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008). The standard software libraries of the ODE numerics can scarcely be used: The individual steps of the MoL-IAFT-FD are permanently linked to the steps of the fully vectorized impedance/admittance and field transformation. In addition, MoL is largely universal and is in principle able to treat almost all possible waveguiding structures in microwave technology and optics. However, the resulting generalized transmission line equations (GTL) can pose different challenges to the efficiency of the finite differences (FD) - depending on the specific waveguiding structure and the excitation. For example, they can be more or less mathematically stiff, (Bronstein et al, 2005), and generally require a particularly high numerical effort or even cause the calculations to be aborted. However, these possible difficulties are by no means the disadvantage of MoL as a method, but a specific feature of the waveguiding structure investigated. In this example, a more efficient solution would be to use a specific (e.g., implicit) method with an optimal choice of the order of accuracy. In order to reduce the numerical effort, it makes sense that the users of the MoL-IAFT-FD should have as many methods as possible to choose from. Initial systematics and recommendations for users were derived (section 4.4.5).

The paper is organized as follows. Section 2 presents the relevant background and shows the two already established context-related methods or types of interpolation. Section 3 represents the proposed extension. The verification of numerical results can be found in section 4. The conclusions are listed in section 5.

2 Background

2.1 Method of Lines

The MoL-IAFT-FD is integrated in the individual steps of the impedance/admittance and field transformation.

Various complex structures, e.g., microwave technology and optics, can be modeled with it, e.g., fiber gratings, (Pregla, 2004), photonic crystals (PC), (Barcz, Helfert and Pregla, 2002), effects of heat propagation, (Conradi, Helfert and Pregla, 2001), and many others (Pregla and Helfert, 2002), (Pregla, 2008). The procedure is as follows: The calculation region is covered with lines. The image of a structure is divided into homogeneous sections in the direction of the analytical solution. The numerical analysis generally consists of two parts:

- Solving wave equations in homogeneous sections
- Field matching at ports between the homogeneous sections

The wave equations are mostly derived with the help of generalized transmission line equations (GTL) (Pregla, 1999), (Pregla, 2002). Then the wave equations or GTL are discretized. Various discretization schemes are available for an efficient analysis (Pregla and Helfert, 2002), (Pregla, 2008), (Greda, 2004). They can be set up in different coordinates, equidistant and non-equidistant, in 2D or 3D (Pregla and Helfert, 2002), (Greda, 2004), (Pregla, 2008). For most practical cases, the discretization of the coordinates perpendicular to the direction of propagation is assumed to be favorable (Vietzorreck, 2001). In the case of 3D, for example, the cross-section of a structure is discretized and the analytical solution is used in the direction of propagation. All details of the discretization and boundary conditions are shown in (Pregla, 2008).

2.2 Generalized transmission line equations

The **g**eneralized **t**ransmission **l**ine equations (GTL) describe the relationship between the transverse components of the electric and the magnetic field. The starting point for deriving the GTL is Maxwell's equations, taking into account the boundary conditions of concrete structures. The detailed representation of the general GTL, corresponding wave equations, the tensor of the material parameters and the normalization of the linear masses, as well as the magnetic field strength is shown in (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008). The following relevant aspects are briefly repeated.

The H- and E field components are discretized on two H and E line systems that are shifted from one another to a discretization distance across the direction of propagation, i.e., z. The discretized field components

are collected in vectors $\widehat{\mathbf{E}}$ and $\widehat{\mathbf{H}}$, or $\widehat{\mathbf{F}} = \left[\widehat{\mathbf{E}}^t, \widehat{\mathbf{H}}^t\right]^t$, corresponding to the spatial distribution of the complex amplitudes of the respective cross section (a step of the FD). The electric and magnetic fields are calculated on two adjacent lines, for details see in (Pregla, 2008), p. 15. The GTL have the general form:

$$\frac{d}{d\overline{u}}\widehat{\mathbf{F}} = \widehat{\boldsymbol{Q}}\widehat{\mathbf{F}} = -\begin{bmatrix} \boldsymbol{S}_{\mathrm{E}} & j\boldsymbol{R}_{\mathrm{H}} \\ j\boldsymbol{R}_{\mathrm{E}} & \boldsymbol{S}_{\mathrm{H}} \end{bmatrix}\widehat{\mathbf{F}} \tag{1}$$

The GTL equation applies to any homogeneous section, e.g., for each step of the FD (section 2.5). The complex matrix $\widehat{\boldsymbol{Q}}$ contains differential operators with corresponding boundary conditions and diagonal matrices of the discretized material parameters. Only lossless isotropic material parameters are considered in the paper. If an isotropy is assumed ($\boldsymbol{S}_{\mathrm{E,H}} = 0$), then the form of the resulting GTL equations

$$\frac{d}{d\overline{u}}\widehat{\mathbf{H}} = -j\widehat{\boldsymbol{R}}_{\mathrm{E}}\widehat{\mathbf{E}} \qquad \frac{d}{d\overline{u}}\widehat{\mathbf{E}} = -j\widehat{\boldsymbol{R}}_{\mathrm{H}}\widehat{\mathbf{H}} \tag{2}$$

resembles the form of the well-known telegrapher's equations from the transmission line theory, (Chen, 2004) ch.V. This analogy is characteristic of the GTL equations and it determines the nature of the underlying calculation processes. It will also be useful in the later discussion about the stability of the calculation. In the case of diagonal material tensors, no losses and lossless boundaries (i.e. Neumann or Dirichlet) the matrices $\widehat{\boldsymbol{R}}_{\mathrm{E,H}}$ are real, symmetrical and in practical applications indefinite. This indefiniteness, as will be discussed later, only rarely appears to disturb the result because of the specific nature of the calculation. The generalized coordinate $u = x, y, z$ is normalized with the free space wave number $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$ according to $\overline{u} = k_0 u$. In addition, the magnetic field components H_u are normalized with the wave impedance $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$ according to $\widetilde{H}_u = \eta_0 H_u$. All further aspects are shown in detail in (Pregla, 2008).

The GTL solution for the whole multi-sectioned device structure is carried out in two procedures: The impedances or admittances are transformed from the output to the input. With the input impedance/admittance, the transverse electric and magnetic field components at the input are obtained. Then these quantities are transformed back to the output. In contrast to the transmission line theory, more than just one mode is considered (the number of modes that are considered is equal to the number of lines). Each of these procedures is carried out for the entire structure in the FD steps. The field transformation serves to determine the fields in each homogeneous section or (which is identical) in each step of the FD. The concept of impedance and admittance matrix transformation is used to analyze such multi-sectioned structures. These matrices are transformed from one side of a section to the other one and from one side of an interface between two sections to the other one. In many cases it is advantageous to perform the impedance/admittance transformation and the field transformation with the help of the matrix parameters $\boldsymbol{V}_{\mathrm{E}}$, $\boldsymbol{Z}_{\mathrm{H}}$, $\boldsymbol{Y}_{\mathrm{E}}$ and $\boldsymbol{V}_{\mathrm{H}}$, see below.

2.2.1 An example: Defect waveguide in a 2D photonic crystal in Cartesian coordinates

Some results will be obtained with this example. It is shown as an application illustration only, without further details. This example relates to a straight defect waveguide in a 2D photonic crystal with hexagonal aligned circular dielectric rods in air, Fig. 1. In this example it seems appropriate to split the GTL into two equations, one for the magnetic and one for the electric field, see below. The content of the corresponding matrices as coefficients of the GTL, $\boldsymbol{R}_{\mathrm{E}}$ and $\boldsymbol{R}_{\mathrm{H}}$, corresponds to (Pregla, 2008): These matrices are built up from diagonal matrices of the material parameters, μ_{uu} and ε_{uu} ($u = x, y, z$), magnetic and dielectric permeability, discretized for a cross section of the structure W on the corresponding step of the FD, and differential operators, $\overline{\boldsymbol{D}}_{\overline{u}}$, which also take corresponding boundary conditions into account. The Neumann boundary conditions for the normal component of the electric field were assumed on the left and right sides of the structure in Fig. 1. However, in the case of the defect waveguide in a photonic crystal and the operating frequency within its bandgap, the wave guide is concentrated on the near area of the defect waveguide. The field strength decreases very quickly on both sides of the waveguide. The effect of the boundary conditions is therefore negligibly small in this case.

The lengths of the (approximately) homogeneous sections in the y-direction are equal to the step of the finite differences. The respective electric and magnetic fields are discretized on two adjacent lines, "$\circ$" and "$\bullet$", that are shifted from one another to a discretization distance across the direction of propagation y. The analytical solution is carried out in the y-direction, whereby the TM and the TE polarization can be analyzed.

TE_y polarization:

E_z, H_x, H_y, $\bullet$-lines, $\mathbf{E} = \mathbf{E}_z$ and $\mathbf{H} = \mathbf{H}_x$, with $\mathbf{H}_y = j\mu_{yy}^{-1}\overline{\boldsymbol{D}}_{\overline{x}}^{\bullet}\mathbf{E}_z$, $\overline{\boldsymbol{D}}_{\overline{z}} = 0$, $\mathbf{E}_x = 0$:

GTL equations: $\frac{d}{d\overline{y}}\widetilde{\mathbf{H}}_x^{\bullet} = -j\boldsymbol{R}_{\mathrm{E}}^{TE_y}\mathbf{E}_z^{\bullet}$ and $\frac{d}{d\overline{y}}\mathbf{E}_z^{\bullet} = -j\boldsymbol{R}_{\mathrm{H}}^{TE_y}\widetilde{\mathbf{H}}_x^{\bullet}$

$$\boldsymbol{R}_{\mathrm{E}}^{\bullet TE_y} = \varepsilon_{zz} - \overline{\boldsymbol{D}}_{\overline{x}}^{\bullet t}\mu_{yy}^{-1}\overline{\boldsymbol{D}}_{\overline{x}}^{\bullet} \quad \boldsymbol{R}_{\mathrm{H}}^{\bullet TE_y} = \mu_{xx} \tag{3}$$

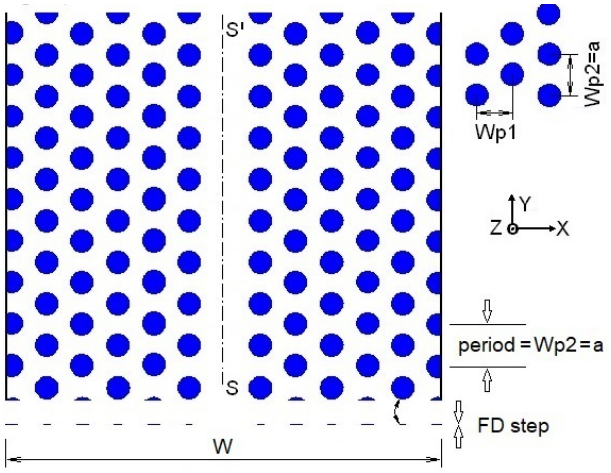


Fig. 1 Defect waveguide in 2D photonic crystal as an example. The wave propagation and the analytical solution are in the y-direction. "W" denotes the width of the structure, 1D discretized cross-section. The structure is shown schematically. It is assumed that the length of the FD step in the y direction is small enough that the section of this length can be assumed to be homogeneous.

TM_y polarization:

H_z, E_x, E_y, $\circ$-lines, $\mathbf{E} = -\mathbf{E}_x$ and $\mathbf{H} = \mathbf{H}_z$. The third field component is obtained from $\mathbf{E}_y = -j\varepsilon_{yy}^{-1}\overline{\boldsymbol{D}}_{\overline{x}}^{\circ}\mathbf{H}_z$, $\overline{\boldsymbol{D}}_{\overline{z}} = 0$, $\mathbf{E}_z = 0$:

GTL equations: $\frac{d}{d\overline{y}}\widetilde{\mathbf{H}}_z^{\circ} = -j\boldsymbol{R}_{\mathrm{E}}^{TM_y}\mathbf{E}_x^{\circ}$ and $-\frac{d}{d\overline{y}}\mathbf{E}_x^{\circ} = -j\boldsymbol{R}_{\mathrm{H}}^{TM_y}\widetilde{\mathbf{H}}_z^{\circ}$

$$\boldsymbol{R}_{\mathrm{E}}^{\circ TM_y} = \varepsilon_{xx} \quad \boldsymbol{R}_{\mathrm{H}}^{\circ TM_y} = \mu_{zz} - \overline{\boldsymbol{D}}_{\overline{x}}^{\circ t}\varepsilon_{yy}^{-1}\overline{\boldsymbol{D}}_{\overline{x}}^{\circ} \tag{4}$$

Note: μ_{xx} is always a appropriate unit matrix. The magnetic field normalized with the wave impedance in vacuum η_0 is marked by the symbol " $\widetilde{\ }$ "; $\widetilde{H}_u = \eta_0 H_u$.

2.3 Impedance/admittance transformation

In the case of a structure composed of different homogeneous sections, the tangential field components must be matched to the transitions. The impedance/admittance transformation is one such matching over the homogeneous sections. The basic idea is the follows. The two-port network parameters of the sections with subports A and B can be calculated from the conditions for open circuit and short circuit set at the output of the structure. This calculation is done step by step, section by section, in the direction from the output of the structure to the input. As a final result, all two-port network parameters of all sections and thus also their respective loads are known. Here is an example of the recursive calculation of the impedance and admittance transformation formulas, respectively:

$$\boldsymbol{Z}_{\mathrm{A}} = \boldsymbol{z}_{11} - \boldsymbol{z}_{12}\left(\boldsymbol{z}_{22} + \boldsymbol{Z}_{\mathrm{B}}\right)^{-1}\boldsymbol{z}_{21} \tag{5}$$

$$\boldsymbol{Y}_{\mathrm{A}} = \boldsymbol{y}_{11} - \boldsymbol{y}_{12}\left(\boldsymbol{y}_{22} + \boldsymbol{Y}_{\mathrm{B}}\right)^{-1}\boldsymbol{y}_{21} \tag{6}$$

$\boldsymbol{Z}_{\mathrm{A,B}}$ and $\boldsymbol{Y}_{\mathrm{A,B}}$ are the impedances and admittances, $\boldsymbol{z}_{\mathrm{ij}}$ and $\boldsymbol{y}_{\mathrm{ij}}$ are the two-port network impedances and admittances, respectively. For the transformation in the opposite direction one obtain:

$$-\boldsymbol{Z}_{\mathrm{B}} = \boldsymbol{z}_{22} - \boldsymbol{z}_{21}\left(\boldsymbol{z}_{11} + (-\boldsymbol{Z}_{\mathrm{A}})\right)^{-1}\boldsymbol{z}_{12} \tag{7}$$

$$-\boldsymbol{Y}_{\mathrm{B}} = \boldsymbol{y}_{22} - \boldsymbol{y}_{21}\left(\boldsymbol{y}_{11} + (-\boldsymbol{Y}_{\mathrm{A}})\right)^{-1}\boldsymbol{y}_{12} \tag{8}$$

An alternative calculation is also used in the paper. The following expressions for a recursive calculation of the input impedance/admittance can easily be derived from the expressions given in (Pregla, 2008), 5.3:

$$\boldsymbol{Z}_{\mathrm{A}} = \left(\boldsymbol{Z}_{\mathrm{H}} + \boldsymbol{V}_{\mathrm{E}}\boldsymbol{Z}_{\mathrm{B}}\right)\left(\boldsymbol{V}_{\mathrm{H}} + \boldsymbol{Y}_{\mathrm{E}}\boldsymbol{Z}_{\mathrm{B}}\right)^{-1} \tag{9}$$

$$\boldsymbol{Y}_{\mathrm{A}} = \left(\boldsymbol{Y}_{\mathrm{E}} + \boldsymbol{V}_{\mathrm{H}}\boldsymbol{Y}_{\mathrm{B}}\right)\left(\boldsymbol{V}_{\mathrm{E}} + \boldsymbol{Z}_{\mathrm{H}}\boldsymbol{Y}_{\mathrm{B}}\right)^{-1} \tag{10}$$

The above expressions apply to a step n of an impedance/admittance transformation. In example, the input impedance of the following section $(n+1)$, the discretized matrix $\boldsymbol{Z}_{\mathrm{B}}$, is related to the input impedance/admittance of the current section $\boldsymbol{Z}_{\mathrm{A}}$. The discretized matrix $\boldsymbol{Z}_0$ is the characteristic impedance of the current homogeneous section. These expressions are similar to the formulas given in (Pregla, 2008) for the mode domain, but applies in this representation to the original domain.

2.4 Field transformation

During field transformation, the electric and magnetic fields are determined using the known impedance and admittances. The field transformation runs section by section in the direction from the input to the output of the structure. The start value of the field at the input of the structure is specified. For example (as in the paper below), it can be a Floquet fundamental mode transformed into the original domain.

This paper focuses on the recursive field transformation using the transmission matrices $\widehat{\boldsymbol{V}}$ in the original domain. The calculation of the field can be done both "forward", from the input to the output of the homogeneous section and "backward", from the output to the input. The side facing the input of the structure is denoted to as (Subport) A, and the side facing the output of the structure as (Subport) B. The corresponding transmission matrices $\widehat{\boldsymbol{V}}_{\mathrm{AB}}$ and $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ are set up for each individual section A-B. The field transformation occurs according to

$$\widehat{\mathbf{F}}_{\mathrm{A}} = \widehat{\boldsymbol{V}}_{\mathrm{AB}} \widehat{\mathbf{F}}_{\mathrm{B}} \tag{11}$$

$$\widehat{\mathbf{F}}_{\mathrm{B}} = \widehat{\boldsymbol{V}}_{\mathrm{BA}} \widehat{\mathbf{F}}_{\mathrm{A}} \tag{12}$$

The use of the four parameters for the field transformation is already specified in (Pregla, 2008):

$$\widehat{\boldsymbol{V}}_{\mathrm{AB}} = \begin{bmatrix} \boldsymbol{V}_{\mathrm{E}} & \boldsymbol{Z}_{\mathrm{H}} \\ \boldsymbol{Y}_{\mathrm{E}} & \boldsymbol{V}_{\mathrm{H}} \end{bmatrix} \qquad \widehat{\boldsymbol{V}}_{\mathrm{BA}} = \widehat{\boldsymbol{V}}_{\mathrm{AB}}^{-1} \tag{13}$$

The matrix $\widehat{\boldsymbol{V}}$ will serve as a container for the new interpolation methods in the paper. Its four components can easily be extracted and used for the impedance/admittance transformation.
The content of the transmission matrices consists of the material parameters and the differential operators with the corresponding boundary conditions. The transmission matrices depend on length of the respective section. All details are listed in (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008).

2.5 Impedance/admittance transformation and field transformation with finite differences

In the case of structures with a complex distribution of the material parameters, e.g., photonic crystals, the impedance/admittance transformation and field transformation can be combined with finite differences, (Helfert and Pregla, 1996), (Pregla, 2008).

The structure is divided into short sections with subports A and B. For sufficiently small distances $\Delta\overline{u} = \tau$ between the subports, the method of finite differences with a corresponding interpolation of the differences can be used. In the past, the two methods of interpolation were built into the GTL solution, the linear and the square, (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008) 2.5.3.

This paper takes up the possibility of expanding the usability of the MoL-IAFT-FD. The approach extends the two previously built-in solution methods in principle to all one-step and multistep methods of various orders of accuracy known in numerical mathematics (Bronstein et al, 2005), (Zeidler, 2004), (Samarski, 1982), (Samarski, 1986). The methods are built into the transmission matrices $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ and/or $\widehat{\boldsymbol{V}}_{\mathrm{AB}}$ for each section of the structure, or in other words, for each step of the FD.

2.6 An uniform use of impedance/admittance and field transformation

In the case of the MoL-IAFT-FD it is possible to calculate the two procedures, the impedance/admittance transformation and field transformation, using the same terms, $\boldsymbol{V}_{\mathrm{E}}$, $\boldsymbol{Z}_{\mathrm{H}}$, $\boldsymbol{Y}_{\mathrm{E}}$ and $\boldsymbol{V}_{\mathrm{H}}$. The FD interpolation is included in the calculation of these four complex matrix parameters and thus influences the quality of both procedures. Before making this connection clear, it should be mentioned that all of the calculations in this paper are performed in the original domain, without a transformation into the eigenmode domain. The possible advantage of working exclusively in the original domain comes from the author's practice: With certain input parameters, working with transfer matrices in the original domain was the only method able to guarantee the computational stability of the

end result (a defect waveguide in 2D photonic crystals with round dielectric rods in air). Possibly the causes of this phenomenon are the object of further research, however, the practical cases have shown that the work exclusively in the original domain is an additional useful instrument of the MoL-IAFT-FD. In addition, the connection between the elements of the impedance/admittance and field transformation enables an uniform treatment in the sense of the FD interpolation. The difference between using these four parameters in the impedance/admittance and field transformation is as follows:

- In the case of the impedance/admittance transformation, these parameters can be used in a recursive calculation of the input impedance, see (9-10).
- In the case of the field transformation, the transfer matrix $\widehat{\boldsymbol{V}}$ is assembled from these parameters, see (13).

2.7 The interpolation methods already integrated in the MoL-IAFT-FD

Let us first briefly introduce the already built-in interpolations, the linear and the quadratic (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008).

2.7.1 The linear interpolation according to (Pregla, 2006-a)

The linear interpolation in (Pregla, 2006-a) corresponds to a more general case of the weighted Euler method (see 3.4) with the parameters $\alpha = 1$ and $\sigma = 0.5$. The accuracy and stability depend on these parameters, and thus also the quality of the solution for specific structures. Further combination of the parameters is considered. At this point, however, the representation according to (Pregla, 2008) follows. Despite the complex content, the general GTL has the form of a partial differential equation of the 1st order

$$\frac{d}{d\overline{u}}\widehat{\mathbf{F}} = \widehat{\boldsymbol{Q}}\widehat{\mathbf{F}} \quad \longleftrightarrow \quad \frac{dy}{d\overline{u}} = f(\overline{u}, y) \tag{14}$$

where the generalized coordinate $u = x, y, z$ is normalized with the wave number k_0, $\overline{u} = k_0 u$. In the following text, all linear dimensions are assumed to be normalized and the overline above them is omitted. It is assumed that the step of the finite differences $\Delta\overline{u} = \tau$ is small enough to guarantee the adequacy of the numerical solution. It is also assumed that the index n is assigned to the input side of the (approximately) homogeneous section, and $n+1$ corresponding to the output side. Then the recursive formula for the numerical solution would be

$$y_{n+1} = y_n + \frac{\tau}{2}(f_{n+1} + f_n) \tag{15}$$

with $n = 1, 2, 3, ..., N$ and N is the number of discretized points for the finite differences. It is applied to the GTL for a homogeneous section:

$$\frac{\widehat{\mathbf{F}}_\mathrm{B} - \widehat{\mathbf{F}}_\mathrm{A}}{\tau} = \frac{1}{2}\left(\widehat{\boldsymbol{Q}}_\mathrm{AB}\widehat{\mathbf{F}}_\mathrm{B} + \widehat{\boldsymbol{Q}}_\mathrm{AB}\widehat{\mathbf{F}}_\mathrm{A}\right) \tag{16}$$

with $\widehat{\boldsymbol{Q}}_\mathrm{AB} = \widehat{\boldsymbol{Q}}(\overline{u}^m)$ and $\overline{u}^m = 0.5(\overline{u}_A + \overline{u}_B)$. The indices "$A$" and "$B$" denote the input and output of the homogeneous section in the direction of propagation. In the following, $\widehat{\mathbf{F}}_\mathrm{A}$ and $\widehat{\mathbf{F}}_\mathrm{B}$ are separated and yield

$$\widehat{\mathbf{F}}_\mathrm{A} = \left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)^{-1}\left(\boldsymbol{I} - \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)\widehat{\mathbf{F}}_\mathrm{B} \qquad \widehat{\boldsymbol{V}}_\mathrm{AB} = \left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)^{-1}\left(\boldsymbol{I} - \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right) \tag{17}$$

or

$$\widehat{\mathbf{F}}_\mathrm{B} = \left(\boldsymbol{I} - \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)^{-1}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)\widehat{\mathbf{F}}_\mathrm{A} \qquad \widehat{\boldsymbol{V}}_\mathrm{BA} = \left(\boldsymbol{I} - \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right)^{-1}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}_\mathrm{AB}\right) \tag{18}$$

The result of this procedure is completely identical to the results of the interpolation of the 1-order in (Pregla, 2008).

2.7.2 The quadratic interpolation according to (Pregla, 2006-b)

The final expressions are shown here. The further details are listed in (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008). The structure and the use of the transmission matrix $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ is shown, for the calculation of the field $\widehat{\mathbf{F}}_{\mathrm{B}}$ from the known field $\widehat{\mathbf{F}}_{\mathrm{A}}$. There is a difference between the first step (or the first homogeneous section of the structure or the first step of the FD) and the subsequent sections.

For $\widehat{\mathbf{F}}_2 = \widehat{\boldsymbol{V}}_1 \widehat{\mathbf{F}}_1$ applies:

$$\widehat{\boldsymbol{V}}_1 = \left[\left(\boldsymbol{I} - \frac{3}{4}\widehat{\boldsymbol{Q}}_1\right) + \frac{1}{8}\widehat{\boldsymbol{Q}}_1\left(\boldsymbol{I} - \frac{3}{8}\widehat{\boldsymbol{Q}}_2\right)^{-1}\left(\boldsymbol{I} + \frac{3}{4}\widehat{\boldsymbol{Q}}_2\right)\right]^{-1}\left[\left(\boldsymbol{I} + \frac{3}{8}\widehat{\boldsymbol{Q}}_1\right) + \frac{1}{64}\widehat{\boldsymbol{Q}}_1\left(\boldsymbol{I} - \frac{3}{8}\widehat{\boldsymbol{Q}}_2\right)^{-1}\widehat{\boldsymbol{Q}}_2\right] \tag{19}$$

and for $\widehat{\mathbf{F}}_{\mathrm{n+1}} = \widehat{\boldsymbol{V}}_{\mathrm{n}} \widehat{\mathbf{F}}_{\mathrm{n}}$ applies:

$$\widehat{\boldsymbol{V}}_{\mathrm{n}} = \left(\boldsymbol{I} - \frac{3}{8}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1}\left(\boldsymbol{I} + \frac{3}{4}\widehat{\boldsymbol{Q}}_{\mathrm{n}} - \frac{1}{8}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\widehat{\boldsymbol{V}}_{\mathrm{n\text{-}1}}^{-1}\right) \qquad n \geq 2 \tag{20}$$

where:

$$\widehat{\boldsymbol{Q}}_{\mathrm{n}} = \Delta\overline{u}\widehat{\boldsymbol{Q}}(\overline{u}_n^m) \qquad \overline{u}_n^m = 0.5(\overline{u}_n + \overline{u}_{n+1}) \qquad \overline{u_1} \equiv \overline{u}_A \tag{21}$$

and $\widehat{\boldsymbol{V}}_{\mathrm{BA}} = \widehat{\boldsymbol{V}}_{\mathrm{n}}$.

The composition of the transmission matrix $\widehat{\boldsymbol{V}}_{\mathrm{AB}}$ is now shown. It should be noted that here, too, when the calculation process moves from the output of the structure towards the input, the calculation of the first step (section) differs from the following steps.

For $\widehat{\mathbf{F}}_{\mathrm{N\text{-}1}} = \widehat{\boldsymbol{V}}_{\mathrm{N}} \widehat{\mathbf{F}}_{\mathrm{N}}$ applies:

$$\widehat{\boldsymbol{V}}_{\mathrm{N}} = \left[\boldsymbol{I} + \frac{3}{4}\widehat{\boldsymbol{Q}}_{\mathrm{N}} - \frac{1}{8}\widehat{\boldsymbol{Q}}_{\mathrm{N}}\left(\boldsymbol{I} + \frac{3}{8}\widehat{\boldsymbol{Q}}_{\mathrm{N\text{-}1}}\right)^{-1}\left(\boldsymbol{I} - \frac{3}{4}\widehat{\boldsymbol{Q}}_{\mathrm{N\text{-}1}}\right)\right]^{-1}\left[\boldsymbol{I} - \frac{3}{8}\widehat{\boldsymbol{Q}}_{\mathrm{N}} + \frac{1}{64}\widehat{\boldsymbol{Q}}_{\mathrm{N}}\left(\boldsymbol{I} + \frac{3}{8}\widehat{\boldsymbol{Q}}_{\mathrm{N\text{-}1}}\right)^{-1}\widehat{\boldsymbol{Q}}_{\mathrm{N\text{-}1}}\right] \tag{22}$$

$$\widehat{\boldsymbol{Q}}_{\mathrm{N}} = \Delta\overline{u}\widehat{\boldsymbol{Q}}(\overline{u}_N^m) \qquad \widehat{\boldsymbol{Q}}_{\mathrm{N\text{-}1}} = \Delta\overline{u}\widehat{\boldsymbol{Q}}(\overline{u}_{N-1}^m) \tag{23}$$

and for $\widehat{\mathbf{F}}_{\mathrm{n\text{-}1}} = \widehat{\boldsymbol{V}}_{\mathrm{n}} \widehat{\mathbf{F}}_{\mathrm{n}}$:

$$\widehat{\boldsymbol{V}}_{\mathrm{n}} = \left(\boldsymbol{I} + \frac{3}{8}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1}\left(\boldsymbol{I} - \frac{3}{4}\widehat{\boldsymbol{Q}}_{\mathrm{n}} + \frac{1}{8}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\widehat{\boldsymbol{V}}_{\mathrm{n+1}}^{-1}\right) \qquad n \leq N-1 \tag{24}$$

Interpreting $\widehat{\mathbf{F}}_{\mathrm{n\text{-}1}}$ as $\widehat{\mathbf{F}}_{\mathrm{A}}$ and $\widehat{\mathbf{F}}_{\mathrm{n}}$ as $\widehat{\mathbf{F}}_{\mathrm{B}}$ from the further sections are calculated using $\widehat{\boldsymbol{V}}_{\mathrm{AB}} = \widehat{\boldsymbol{V}}_{\mathrm{n}}$. For further details, see e.g., (Pregla, 2008).

3 An expansion of the GTL solutions with further one-step procedures

3.1 General considerations

The quality of the MoL-IAFT-FD essentially depends on the effectiveness of the used interpolation of differences. However, the quality of the one or another kind of interpolation depends on the concrete behavior of the solution sought, e.g., (Bultheel and Cools, 2010), (Curtiss and Hirschfelder, 1952), comp. (Spiller et al., 2019), i.e., on specific applications, or, on each specific waveguiding structure:

- on the spatial configuration and specific values of the material parameters
- on the order of accuracy of the method and its approach: explicitly or implicitly.

The way to increase the usability of MoL-IAFT-FD is to incorporate further methods from numerical mathematics into the individual steps of the GTL solution. The methods should be interchangeable with as little effort as possible, and the user can choose to use them for various specific applications. Two procedures for the numerical solution of ordinary differential equations comes into consideration: one-step and multi-step methods of some order of accuracy (Bronstein et al, 2005), (Zeidler, 2004), (Samarski, 1982), (Samarski, 1986), (Hairer et al, 2007), (Bultheel and Cools, 2010), (Hairer and Wanner, 1996). The one-step method needs only one of the previous values y_n to calculate the function value at the next step y_{n+1} ($n = 1, 2, 3, ..., N$ and N is the number of discretized points). The multistep procedures need several preceding values, e.g., y_n and y_{n-1}. The assignment to one or the other method will be of particular interest to us with regard to accuracy and stability: For certain complex structures such as Bragg gratings, photonic crystals, etc., the accuracy and stability can be a challenge.

For the improvement of accuracy, — while preserving numerical stability, — the higher-order methods can be used. However, the aspect of numerical stability limits the choice of methods: The Second Dahlquist Barrier states that the consistency order of an A-stable multistep method can be at most two (Dahlquist, 1963). But it does not apply to the one-step methods such as Runge-Kutta. Moreover, the Daniel-Moore conjecture follows, that the implicit Runge-Kutta methods can be of any order of accuracy (Hairer et al, 2007), (Bultheel and Cools, 2010), (Hairer and Wanner, 1996). Therefore, a wide choice of one-step methods of the higher-order appears to be useful.

The two interpolations that have been build in into the MoL-IAFT-FD in the past are the linear and the square (Pregla, 2006-a), (Pregla, 2006-b), (Pregla, 2008). They are the representatives of the one-step and multistep methods, respectively. In principle, the multi-step methods can be built into the MoL-IAFT-FD the same way as the one-step methods. To keep the papers a little simpler, let's turn to the one-step methods.

However, higher-order methods are not necessarily more accurate, - because of stability problems or discretization errors (Bultheel and Cools, 2010). In some cases, an interpolation of the lower order may be even advantageous (Spiller et al., 2019). This consideration also speaks in favor of a broader choice, including the lower orders.

Another numerical aspect can also be relevant for complex structures: the mathematical stiffness of the GTL. In general, differential equations are mathematically "stiff" if they contain some constructs or parameters that cause rapid variations in the solutions. Here, too, the example of certain photonic crystals would apply. It is generally difficult to integrate the "stiff" equations by ordinary numerical methods. Small errors may rapidly accumulate (Bronstein et al, 2005), (Zeidler, 2004), (Curtiss and Hirschfelder, 1952), (Hairer and Wanner, 1996). In many cases, the implicit methods that are more tolerant of the stiffness can be considered.

The existing experience in industry and research also speaks in favor of a broader choice of methods: The standard libraries of many universal mathematical software products have a variable-step, variable-order special solvers of orders 1–5. A solver is a piece of software and can perform basic operations. However, the use of these solvers in the MoL-IAFT-FD is not easily possible: The steps of the FD solution are integrated into the context of the impedance/admittance transformation and field transformation, (see section 2.3). From these considerations, it follows: An appropriate choice of the appropriate type of interpolation being MoL-IAFT-FD should be an effective aid. For the purpose of applicability, the MoL-IAFT-FD set of instruments should have the option of a broader selection of integrated methods. These methods should be easily interchangeable or as portable as possible for the user.

This paper takes up the possibility of incorporating practically every numerical method from the class of one-step and multi-step methods (Bronstein et al, 2005), (Zeidler, 2004), (Samarski, 1982), (Samarski, 1986) into the GTL solution. This is done in a uniform way by modifying transmission matrices. These can be methods of the higher-orders, explicit and implicit. In this paper, the one-step methods are built into the GTL solution. The installation of some additional multistep methods would, however, be possible in the same way.
Based on the considerations, it will focus on the one-step methods in a uniform way.

- The transmission matrices $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ and/or $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ serve as "containers" for various built-in methods and as the "end product" of our considerations
- An individually assembled transmission matrix is required for each step of the analysis procedure
- The calculation of the field with the aid of transmission matrices is in principle equivalent to the recursive procedures with the z- or y-parameters in the transformed domains (modes or Floquet mode domains). However, the use of transmission matrices appears to be the easiest way to incorporate other different methods.

Seven further methods are built in: Both explicit and implicit Euler methods of order 1 of accuracy, Euler method with weighting as a general case for the Euler explicit/implicit approach, two kinds of the Runge-Kutta methods of order 2 (RK2-I and RK2-II), classical explicit Runge-Kutta method of order 4 (RK4) and implicit Gauss-Runge-Kutta method of order 4 (GRK4), (Hoellig, 2011), (Hoellig, 1998). These are tested on simple and "difficult" structures, and the quality of the results is compared with one another and with the already known solutions (see section 4).

3.2 Euler method explicit

The relation can be written as $\frac{y_{n+1}-y_n}{\tau} = \alpha f_n \implies y_{n+1} = y_n + \alpha\tau f_n$ (Samarski, 1986). This is used to solve the GTL:

$$\widehat{\mathbf{F}}_{\mathrm{B}} = \widehat{\mathbf{F}}_{\mathrm{A}} + \alpha\tau\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} = \left(\boldsymbol{I} + \alpha\tau\widehat{\boldsymbol{Q}}\right)\widehat{\mathbf{F}}_{\mathrm{A}} \tag{25}$$

It is formally

$$\widehat{\boldsymbol{Q}}_{\mathrm{n}} = \Delta\overline{u}\widehat{\boldsymbol{Q}}(\overline{u}_n^m) \qquad \overline{u}_n^m = 0.5(\overline{u}_n + \overline{u}_{n+1}) \qquad \overline{u_1} \equiv \overline{u}_A \tag{26}$$

although applies to the entire homogeneous section.

$$\widehat{\boldsymbol{V}}_{\mathrm{AB}} = \left(\boldsymbol{I} + \alpha\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1} \qquad \widehat{\boldsymbol{V}}_{\mathrm{BA}} = \left(\boldsymbol{I} + \alpha\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) \tag{27}$$

The subscript "AB" shows that the matrices $\widehat{\boldsymbol{V}}_{\mathrm{AB}}$ and $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ only applies to the respective section.

3.3 Euler method implicit

The implicit procedure $\frac{y_{n+1}-y_n}{\tau} = \alpha f_{n+1} \quad \Longrightarrow \quad y_{n+1} = y_n + \alpha\tau f_{n+1}$ is assumed, (Samarski, 1986). α is a selectable dimensionless parameter. For a better comparison of the methods, $\alpha = 1/2$ is chosen. This results in terms of (14) in

$$\widehat{\mathbf{F}}_{\mathrm{A}} = \widehat{\mathbf{F}}_{\mathrm{B}} - \alpha\tau\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{B}} = \left(\boldsymbol{I} - \alpha\tau\widehat{\boldsymbol{Q}}\right)\widehat{\mathbf{F}}_{\mathrm{B}} \tag{28}$$

It all results in

$$\widehat{\boldsymbol{V}}_{\mathrm{AB}} = \left(\boldsymbol{I} - \alpha\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) \qquad \widehat{\boldsymbol{V}}_{\mathrm{BA}} = \left(\boldsymbol{I} - \alpha\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1} \tag{29}$$

3.4 Euler method with weighting (Euler-W)

The procedure with weighting is a general case (also for the linear Interpolation in (Pregla, 2006-a), see 2.7.1) in which one can use the parameter σ to vary the properties of the method between explicit and implicit, (Samarski, 1986):

$$\frac{y_{n+1}-y_n}{\tau} = \alpha\left(\sigma f_{n+1} + (1-\sigma) f_n\right) \quad \Longrightarrow \quad y_{n+1} = y_n + \alpha\tau\left(\sigma f_{n+1} + (1-\sigma) f_n\right) \tag{30}$$

It is used to solve the GTL:

$$\frac{\widehat{\mathbf{F}}_{\mathrm{B}} - \widehat{\mathbf{F}}_{\mathrm{A}}}{\tau} = \alpha\sigma\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{B}} + \alpha(1-\sigma)\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} \tag{31}$$

It results in

$$\widehat{\mathbf{F}}_{\mathrm{B}} = \left(\boldsymbol{I} - \alpha\sigma\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1}\left(\boldsymbol{I} + \alpha(1-\sigma)\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)\widehat{\mathbf{F}}_{\mathrm{A}} \tag{32}$$

or

$$\widehat{\boldsymbol{V}}_{\mathrm{BA}} = \left(\boldsymbol{I} - \alpha\sigma\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right)^{-1}\left(\boldsymbol{I} + \alpha(1-\sigma)\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) \qquad \widehat{\boldsymbol{V}}_{\mathrm{AB}} = \widehat{\boldsymbol{V}}_{\mathrm{BA}}^{-1} \tag{33}$$

$\alpha = 1$ and $\sigma = 0.8$ is selected for comparability with the other methods.

3.5 Runge–Kutta methods

The possible Runge-Kutta methods of the second and fourth order of accuracy, (Samarski, 1982), (Samarski, 1986), are built-in into all the solution steps of the GTL.

3.5.1 Second-order methods (RK2)

In general, the calculations of the 2nd order method are performed in two steps. First, the intermediate result $\overline{y}_n$ is found according to the Euler scheme with the step length $\alpha\tau$:

$$\overline{y}_n = y_n + \alpha\tau f(\overline{u}_n, y_n) \tag{34}$$

In the second step y_{n+1} is found:

$$y_{n+1} = y_n + \tau(1-\sigma) f(\overline{u}_n, y_n) + \sigma\tau f(\overline{u}_n + \alpha\tau, \overline{y}_n) \tag{35}$$

where $\alpha > 0$ and $\sigma > 0$ are the selectable parameters. After excluding $\overline{y}_n$, the Runge-Kutta scheme of the 2nd order results in

$$\frac{y_{n+1}-y_n}{\tau} = (1-\sigma) f(\overline{u}_n, y_n) + \sigma f(\overline{u}_n + \alpha\tau, y_n + \alpha\tau f(\overline{u}_n, y_n)) \tag{36}$$

The order of accuracy depends on the parameters α and τ. Satisfying the condition $\sigma\alpha = 1/2$ results in a schemes family (36) of the 2nd order of accuracy (Samarski, 1982), (Samarski, 1986).

3.5.2 The predictor-corrector scheme of the 2nd order, variant I (RK2-I)

The variant $\alpha = 1/2$, $\sigma = 1$ is being considered. This is the well-known scheme "predictor-corrector" (Samarski, 1982), (Samarski, 1986). It can be represented in the form

$$\overline{y}_n = y_n + \frac{\tau}{2} f(\overline{u}_n, y_n), \qquad y_{n+1} = y_n + \tau f\left(\overline{u}_n + \frac{\tau}{2}, \overline{y}_n\right) \tag{37}$$

or, after excluding $\overline{y}_n$,

$$\frac{y_{n+1} - y_n}{\tau} = f\left[\overline{u}_n + \frac{\tau}{2}, y_n + \frac{\tau}{2} f\left(\overline{u}_n, y_n\right)\right] \tag{38}$$

After applying (35) to the parameters of the GTL, the result is

$$\frac{\widehat{\mathbf{F}}_{\mathrm{B}} - \widehat{\mathbf{F}}_{\mathrm{A}}}{\tau} = \widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_{\mathrm{A}} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}}\right) \tag{39}$$

$$\widehat{\mathbf{F}}_{\mathrm{B}} = \left[\widehat{\boldsymbol{Q}}_{\mathrm{n}}\left(\boldsymbol{I} + \frac{1}{2}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) + \boldsymbol{I}\right]\widehat{\mathbf{F}}_{\mathrm{A}} \tag{40}$$

$$\widehat{\boldsymbol{V}}_{\mathrm{BA}} = \widehat{\boldsymbol{Q}}_{\mathrm{n}}\left(\boldsymbol{I} + \frac{1}{2}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) + \boldsymbol{I} \qquad \widehat{\boldsymbol{V}}_{\mathrm{AB}} = \widehat{\boldsymbol{V}}_{\mathrm{BA}}^{-1} \tag{41}$$

3.5.3 The predictor-corrector scheme of the 2nd order, variant II (RK2-II)

$\alpha = 1$ and $\sigma = 1/2$: This scheme can also be interpreted as a "predictor-corrector" (Samarski, 1982), (Samarski, 1986): The Euler scheme with the step τ (predictor)

$$\overline{y}_n = y_n + \tau f(\overline{u}_n, y_n) \tag{42}$$

is calculated first, then the scheme with half the sum (corrector):

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\left[f\left(\overline{u}_n, y_n\right) + f\left(\overline{u}_{n+1}, \overline{y}_n\right)\right] \tag{43}$$

After excluding $\overline{y}_n$:

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{2}\left[f\left(\overline{u}_n, y_n\right) + f\left(\overline{u}_{n+1}, y_n + \tau f\left(\overline{u}_n, y_n\right)\right)\right] \tag{44}$$

After applying (44) to the parameters of the GTL, the result is

$$\frac{\widehat{\mathbf{F}}_{\mathrm{B}} - \widehat{\mathbf{F}}_{\mathrm{A}}}{\tau} = \frac{1}{2}\left[\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} + \widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_{\mathrm{A}} + \tau\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}}\right)\right] \tag{45}$$

$$\widehat{\mathbf{F}}_{\mathrm{B}} = \left[\frac{1}{2}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\left(2\boldsymbol{I} + \widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) + \boldsymbol{I}\right]\widehat{\mathbf{F}}_{\mathrm{A}} \tag{46}$$

that results in

$$\widehat{\boldsymbol{V}}_{\mathrm{BA}} = \frac{1}{2}\widehat{\boldsymbol{Q}}_{\mathrm{n}}\left(2\boldsymbol{I} + \widehat{\boldsymbol{Q}}_{\mathrm{n}}\right) + \boldsymbol{I} \qquad \widehat{\boldsymbol{V}}_{\mathrm{AB}} = \widehat{\boldsymbol{V}}_{\mathrm{BA}}^{-1} \tag{47}$$

3.5.4 The 4th order Runge-Kutta method (RK4)

$$\frac{y_{n+1} - y_n}{\tau} = \frac{1}{6}\left(K_1 + 2K_2 + 2K_3 + K_4\right) \tag{48}$$

where, in according to (Samarski, 1982), (Samarski, 1986),

$$K_1 = f\left(\overline{u}_n, y_n\right) \Rightarrow \widehat{\boldsymbol{K}}_1 = \widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} \tag{49}$$

$$K_2 = f\left(\overline{u}_n + \frac{\tau}{2}, y_n + \frac{\tau}{2}K_1\right) \Rightarrow \widehat{\boldsymbol{K}}_2 = \widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_{\mathrm{A}} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}}\right) = \left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\right)\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} \tag{50}$$

$$K_3 = f\left(\overline{u}_n + \frac{\tau}{2}, y_n + \frac{\tau}{2}K_2\right) \Rightarrow \widehat{\boldsymbol{K}}_3 = \widehat{\boldsymbol{Q}}\left[\widehat{\mathbf{F}}_{\mathrm{A}} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_{\mathrm{A}} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}}\right)\right] = \left[\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\right)\right]\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_{\mathrm{A}} \tag{51}$$

$$K_4 = f\left(\overline{u}_n + \tau, y_n + \tau K_3\right) \Rightarrow \widehat{\boldsymbol{K}}_4 = \widehat{\boldsymbol{Q}}\left\{\widehat{\mathbf{F}}_\mathrm{A} + \tau\left[\widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_\mathrm{A} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_\mathrm{A} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_\mathrm{A}\right)\right)\right]\right\} = \\ = \left\{\boldsymbol{I} + \tau\widehat{\boldsymbol{Q}}\left[\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{Q}}\right)\right]\right\}\widehat{\boldsymbol{Q}}\widehat{\mathbf{F}}_\mathrm{A} \quad (52)$$

The coefficients are considered without a field $\widehat{\mathbf{F}}_\mathrm{A}$, where: $\widehat{\boldsymbol{K}}_{1,2,3,4} = \widehat{\boldsymbol{K}}^\mathrm{w}_{1,2,3,4}\widehat{\mathbf{F}}_\mathrm{A}$. It results in

$$\widehat{\boldsymbol{K}}^\mathrm{w}_1 = \widehat{\boldsymbol{Q}} \quad (53)$$

$$\widehat{\boldsymbol{K}}^\mathrm{w}_2 = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{K}}^\mathrm{w}_1\right) \quad (54)$$

$$\widehat{\boldsymbol{K}}^\mathrm{w}_3 = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{\tau}{2}\widehat{\boldsymbol{K}}^\mathrm{w}_2\right) \quad (55)$$

$$\widehat{\boldsymbol{K}}^\mathrm{w}_4 = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \tau\widehat{\boldsymbol{K}}^\mathrm{w}_3\right) \quad (56)$$

$$y_{n+1} = y_n + \frac{\tau}{6}\left(K_1 + 2K_2 + 2K_3 + K_4\right) \quad \Rightarrow \quad \widehat{\mathbf{F}}_\mathrm{B} = \widehat{\mathbf{F}}_\mathrm{A} + \frac{\tau}{6}\left(\widehat{\boldsymbol{K}}^\mathrm{w}_1 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_2 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_3 + \widehat{\boldsymbol{K}}^\mathrm{w}_4\right)\widehat{\mathbf{F}}_\mathrm{A} \\ \widehat{\mathbf{F}}_\mathrm{B} = \left[\boldsymbol{I} + \frac{\tau}{6}\left(\widehat{\boldsymbol{K}}^\mathrm{w}_1 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_2 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_3 + \widehat{\boldsymbol{K}}^\mathrm{w}_4\right)\right]\widehat{\mathbf{F}}_\mathrm{A} \quad (57)$$

$$\widehat{\boldsymbol{V}}_\mathrm{BA} = \boldsymbol{I} + \frac{\tau}{6}\left(\widehat{\boldsymbol{K}}^\mathrm{w}_1 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_2 + 2\widehat{\boldsymbol{K}}^\mathrm{w}_3 + \widehat{\boldsymbol{K}}^\mathrm{w}_4\right), \qquad \widehat{\boldsymbol{V}}_\mathrm{AB} = \widehat{\boldsymbol{V}}^{-1}_\mathrm{BA} \quad (58)$$

3.5.5 The 4th order Gauss-Runge-Kutta implicit method (GRK4)

The implicit method, according to Gauss-Runge-Kutta, (Seyrich, 2016), (Grothmann, 2012), (Grothmann, 2015), (Hoellig, 2011), (Hoellig, 1998), (Pulch, 2020) has the order of convergence 4, the order of accuracy 5 and is suitable for stiff GTL. The predictor is shown first:

$$\frac{y_{n+1} - y_n}{\tau} = \frac{K_1 + K_2}{2} \quad (59)$$

where K_1 and K_2 are given implicitly:

$$K_1 = f\left(\overline{u}_n + \frac{3-\sqrt{3}}{6}\tau, y_n + \frac{1}{4}\tau K_1 + \frac{3-2\sqrt{3}}{12}\tau K_2\right) \quad (60)$$

$$K_2 = f\left(\overline{u}_n + \frac{3+\sqrt{3}}{6}\tau, y_n + \frac{3+2\sqrt{3}}{12}\tau K_1 + \frac{1}{4}\tau K_2\right) \quad (61)$$

The expressions (60) and (61) therefore require an iterative procedure (corrector steps) for K_1 and K_2. These iterations are carried out on each step n of the FD ($n = 1,2,3,...,N$). ${K_1}^{(0)} = f(\overline{u}_n, y_n)$ and ${K_2}^{(0)} = K_1$ can serve as starting values (predictor step). The criterion for the end of the iterations depends on the convergence of the result: The iterations are ended, for example, as soon as the difference between two successive improvements is within the specified tolerance limits.

Next, this consideration is used to solve the GTL with MoL-IAFT-FD. For the predictor

$$\frac{\widehat{\mathbf{F}}_\mathrm{B} - \widehat{\mathbf{F}}_\mathrm{A}}{\tau} = \frac{\widehat{\boldsymbol{K}}_1 + \widehat{\boldsymbol{K}}_2}{2} \quad (62)$$

K_1 and K_2 are put together according to (60) and (61):

$$\widehat{\boldsymbol{K}}_1 = \widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_\mathrm{A} + \frac{1}{4}\tau\widehat{\boldsymbol{K}}_1 + \frac{3-2\sqrt{3}}{12}\tau\widehat{\boldsymbol{K}}_2\right) \quad (63)$$

$$\widehat{\boldsymbol{K}}_2 = \widehat{\boldsymbol{Q}}\left(\widehat{\mathbf{F}}_\mathrm{A} + \frac{3+2\sqrt{3}}{12}\tau\widehat{\boldsymbol{K}}_1 + \frac{1}{4}\tau\widehat{\boldsymbol{K}}_2\right) \quad (64)$$

In order to determine the transmission matrices, the field $\widehat{\mathbf{F}}_\mathrm{A}$ is separated from (63) and (64):

$$\widehat{\boldsymbol{K}}_1 = \widehat{\boldsymbol{K}}^{(\mathrm{w})}_1\widehat{\mathbf{F}}_\mathrm{A} \qquad \widehat{\boldsymbol{K}}_2 = \widehat{\boldsymbol{K}}^{(\mathrm{w})}_2\widehat{\mathbf{F}}_\mathrm{A} \quad (65)$$

$$\widehat{\mathbf{F}}_\mathrm{B} = \left[\boldsymbol{I} + \frac{\tau}{2}\left(\widehat{\boldsymbol{K}}^{(\mathrm{w})}_1 + \widehat{\boldsymbol{K}}^{(\mathrm{w})}_2\right)\right]\widehat{\mathbf{F}}_\mathrm{A} \quad (66)$$

The starting values for the predictor step are

$$\widehat{\boldsymbol{K}}_1^{(\mathrm{w}0)} = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{1}{4}\tau\widehat{\boldsymbol{Q}} + \frac{3-2\sqrt{3}}{12}\tau\widehat{\boldsymbol{Q}}\right) \qquad \widehat{\boldsymbol{K}}_2^{(\mathrm{w}0)} = \widehat{\boldsymbol{K}}_1^{(\mathrm{w}0)} \tag{67}$$

and the iterations are performed with

$$\widehat{\boldsymbol{K}}_1^{(\mathrm{w})} = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{1}{4}\tau\widehat{\boldsymbol{K}}_1^{(\mathrm{w})} + \frac{3-2\sqrt{3}}{12}\tau\widehat{\boldsymbol{K}}_2^{(\mathrm{w})}\right) \tag{68}$$

$$\widehat{\boldsymbol{K}}_2^{(\mathrm{w})} = \widehat{\boldsymbol{Q}}\left(\boldsymbol{I} + \frac{3+2\sqrt{3}}{12}\tau\widehat{\boldsymbol{K}}_1^{(\mathrm{w})} + \frac{1}{4}\tau\widehat{\boldsymbol{K}}_2^{(\mathrm{w})}\right) \tag{69}$$

Thus the transmission matrices are

$$\widehat{\boldsymbol{V}}_{\mathrm{BA}} = \boldsymbol{I} + \frac{\tau}{2}\left(\widehat{\boldsymbol{K}}_1^{(\mathrm{w})} + \widehat{\boldsymbol{K}}_2^{(\mathrm{w})}\right) \qquad \widehat{\boldsymbol{V}}_{\mathrm{AB}} = \widehat{\boldsymbol{V}}_{\mathrm{BA}}^{-1} \tag{70}$$

As a summary for the GRK4: The procedure for calculating for each section n is as follows:

1. Assembling the matrix $\widehat{\boldsymbol{Q}}$
2. Setting of the start values $\widehat{\boldsymbol{K}}_1^{(\mathrm{w}0)}$ and $\widehat{\boldsymbol{K}}_2^{(\mathrm{w}0)}$ and an iterative improvement of the values $\widehat{\boldsymbol{K}}_1^{(\mathrm{w})}$ and $\widehat{\boldsymbol{K}}_2^{(\mathrm{w})}$
3. Assembling the transmission matrices $\widehat{\boldsymbol{V}}_{\mathrm{BA}}$ and/or $\widehat{\boldsymbol{V}}_{\mathrm{AB}}$

4 Verification of results and discussion

4.1 General considerations and test structures

The general properties of the one-step and multistep methods are well researched. It is therefore focused on their application to MoL and complex waveguiding structures. Multimode waveguides can have very complex shapes of field distributions as a result of a superposition of several eigenmodes. Although defect waveguides in photonic crystals can serve as an example, however, even simple structures can exhibit the most complex spatial field distributions if they can carry a large number of modes and are excited accordingly. Such solutions can be a challenge for numerical methods. In the paper, fields that are as complex as possible should be deliberately selected as test objects in order to create difficult test conditions for the methods and to observe their behavior. On the other hand, such a choice is problematic because a spatial superposition of multiple modes is difficult to predict. Thus, even small changes in the phase relationships of the individual modes can cause a significant change in the shape of the resulting field distribution. That is why it is practically difficult to synthesize a suitable multimode excitation, so that the resulting superposition gives at least some test forms of the field distribution, which are difficult for the methods. Although such a synthesis would be possible, the corresponding effort does not appear to be appropriate to the purpose of the paper.Therefore, it seems the only way to at least test the methods exploratively and to draw at least the most basic conclusions from them. It is therefore hypothesized - although not always true - that the occurrence of complex spatial field distributions required for the tests is more likely for structures with complex material parameter distributions (and, e.g., outside the sub-wavelength range). In other words, in the absence of synthesis-based information to generate certain complex spatial field distributions, one can generate complex spatial distributions of the material parameters as a minimum basis. Farther, it is expected that abrupt transitions of the dielectric material parameters and the resulting reflections will add to the complexity of the resulting field distributions. And although this effect can only be controlled to a limited extent, it provides an opportunity to create appropriate test structures by means of deliberately introduced abrupt transitions with suitable contrasts. However, the number of such abrupt transitions should not be too large, - otherwise numerous coexisting modes and reflections will statistically cancel each other out, and the field distributions will become smoother. The requirements for the initial tests follow from this:

- A test structure should be able to carry a sufficiently large number of modes for the working wavelength, e.g., have a sufficiently large aperture and suitable boundary conditions.
- A test excitation shall consist of several modes to allow at least some resulting field distributions that can be challenging for a numerical solution.

In doing so, it is inevitably accepted that the test coverage of such an exploratory procedure is minimal and incomplete. On the other hand, this test procedure appears to be more effective than a testing on simple field distributions, and, therefore most efficient for the purpose of this paper. Accordingly, three multimode test structures are considered: one with a distinctly smooth spatial distribution of the dielectric permittivity in the longitudinal direction ("cosine profile", Fig. 2) and one with two abrupt transitions: A symmetrical rectangular distribution of permittivity ("pulse profile", Fig. 4). A defect waveguide in a photonic crystal (PC) is to serve as the third test structure. This structure is intended to be operated within its bandgap. The guided wave in a cross-section presents a relatively sharp-edged Gaussian curve with possible side lobes that may also be sharp-edged ("PC profile", Fig. 9).

These three test structures appear to cover the three test conditions:

1. Rather smooth field distributions ("cosine profile").
2. Complex Interference-related field distributions with periodically distributed steep portions ("pulse profile").
3. A somewhat smoother field distribution, but with some strong steep parts in places (wave guidance in a photonic crystal).

As shown below (section 4.4.3 and Fig. 2-13), these three test conditions seem allow some basic conclusions despite the incompleteness of the tests (section 4.4.5).

4.2 On the stability, convergence and consistency

4.2.1 Expediency and approach

The matrix $\boldsymbol{Q}$ (or the matrices $\boldsymbol{R}_{\mathrm{E}}$ and $\boldsymbol{R}_{\mathrm{H}}$) can be no longer semidefinite depending on the spatial discretization. How is stability ensured - and, therefore, also convergence?
As a first step, it seems appropriate to discuss a definition suitable for the engineering of the MoL applications. The following properties of a numerical algorithm in the sense of (Bronstein et al, 2005) and (Zeidler, 2004) are considered:

- Stability: How much does the result of the numerical algorithm deviate if the input is disturbed.
- Convergence: How well does the step size dependent algorithm approximate the exact solution. "A method is convergent if and only if it is stable" (Lax-Richtmyer theorem).
- Consistency: How small is the step size dependent on the local truncation error (at each step).

When analyzing the results, the concept of condition - as a property of an original problem - should also be taken into account:

- Condition: How much does the exact solution of the original problem deviate if its input is disturbed?

To assure numerical stability, additional information about the underlying process should be used: The impedance/admittance transformation. This transformation is inherently stable, which is briefly shown below. All investigated methods are based on the impedance/admittance transformation usually show convergence and consistency. The behavior of the impedance/admittance transformation is considered in two cases:

- Very long steps, $\Delta\overline{u} \to \infty$
- Very short steps, $\Delta\overline{u} \to 0$

Next, the following aspects will be discussed:

- Influence of the mathematical stiffness of the GTL, (section 4.2.6)
- Influence of the waveguide specifics on the condition of corresponding models and their solution in practice, (section 4.2.7)

4.2.2 Very long and very short steps

In order to investigate the characteristic features of the methods with regard to the MoL, the fineness of the discretization was investigated in a very wide range (convergence curves Fig. 3, 5-8, 10-13). The coarsest discretization (the left region of the convergence curves Fig. 3, 5-8, 10-13) should show the stability at the limit of the spatial Nyquist–Shannon sampling theorem and thus at the limit of meaningful application. The finest discretization (the right region of the convergence curves Fig. 3, 5-8, 10-13) should enable the estimation of the convergence value, but also show a possible divergence, e.g. due to the accumulation of the rounding error $O(1/\Delta\overline{u})$, (Bronstein et al, 2005), (Zeidler, 2004).

4.2.3 Two limit values of the step length

The impedance/admittance transformation concept is numerically stable and gives correct results for every length of the sections $\Delta\overline{u}$. This fact is based on a direct relation to the transmission line theory, (Chen, 2004) ch.V, which in principle provides exact analytical solutions. Also, in general, the algorithm of the impedance/admittance transformation can be understood as generalized transmission line theory, (Pregla, 2008). Thus, the GTL equations, e.g., (2), show the analogy between the MoL and the well-known telegraph equations. Lines used in the MoL can also be represented as transmission lines. If one has only one mode, the equations (5-8) or (9-10) reduce to to the well-known impedance/admittance transformation formulas from the transmission line theory. Hence, as shown below, the following applies: As long as individual sections (as steps of the FD) are calculated according to the rules of the transmission line theory, the impedance/admittance transformation remains for $\Delta\overline{u} \to \infty$ and $\Delta\overline{u} \to 0$ formally exact, (Pregla, 2008) (e.g., ch. 2). In other words: As long as the two-port parameters of the last section at the output of the structure are determined by short and open circuiting, - one can assume an exact impedance/admittance transformation. Of course, this accuracy relates to the formally calculated mathematical model, which does not necessarily correspond to the original, e.g., because of a too coarse discretization. It is the responsibility of the user to determine whether the discretization is sufficiently dense. If the sampling theorem is disregarded, the results will be incorrect, which, however, has to do with an incorrect application of the model. But even in this case, the impedance/admittance transformation remains formally exact and stable.

In this way, it can be shown that increasing or decreasing the step length does not lead to any formal instability of the end results: the impedance/admittance transformation formulas (5-8) or (9-10) are numerically stable for very long sections (or very thick layers) and for very short sections (or very thin layers), (Pregla, 2008), 2.5.2.1. This applies both to the MoL theory according to (Pregla, 2008) and to the underlying transmission line theory for individual modes. This can be confirmed by checking the corresponding limit cases in both theories. The results are as follows:
$\Delta\overline{u} \to \infty$:
With increasing section thickness $\Delta\overline{u}$, one obtains:

$$\lim_{\Delta\overline{u}\to\infty} \boldsymbol{Z}_{\mathrm{A}} = \boldsymbol{Z}_0 \qquad \lim_{\Delta\overline{u}\to\infty} \boldsymbol{Y}_{\mathrm{A}} = \boldsymbol{Y}_0 \tag{71}$$

The quantities $\boldsymbol{Z}_0$ and $\boldsymbol{Y}_0 = \boldsymbol{Z}_0^{-1}$ are the wave impedances and admittances for each individual section (or step of the FD), respectively.
$\Delta\overline{u} \to 0$:
For very short sections (or very thin layers), the impedance (admittance) $\boldsymbol{Z}_{\mathrm{A}}$ ($\boldsymbol{Y}_{\mathrm{A}}$) approaches $\boldsymbol{Z}_{\mathrm{B}}$ ($\boldsymbol{Y}_{\mathrm{B}}$):

$$\lim_{\Delta\overline{u}\to 0} \boldsymbol{Z}_{\mathrm{A}} = \boldsymbol{Z}_{\mathrm{B}} \qquad \lim_{\Delta\overline{u}\to 0} \boldsymbol{Y}_{\mathrm{A}} = \boldsymbol{Y}_{\mathrm{B}} \tag{72}$$

It should be emphasized again: The method is correct and stable with regard to the formally calculated model, but the user should ensure that the formal model corresponds to the original structure, e.g., in terms of the sampling theorems.

4.2.4 The behavior in the region of a coarse discretization

It is a very low spatial sampling rate of the material parameters at the limit of the sampling theorem. This working region of coarse discretization is only used in practice for special cases when the user is forced to do so by the specifics of the task, e.g., when a large numerical effort has to be avoided. Much more, however, this work region in the paper will serve the goal of a hard test and the comparability between the methods.

The paper starts with such a coarse discretization that a correct application of the methods just seems possible (the left region of the convergence curves Fig. 3, 5-8, 10-13). An even coarser discretization would already cause an “instability” in the sense of the complete loss of the adequacy of the model, (Bronstein et al, 2005), (Zeidler, 2004). Here, too, it should be emphasized that the MoL remains formally stable and “accurate” even with a very coarse discretization, i.e., even if the resulting model deviates completely from the user’s original goal.

A very rough discretization at the limit of the sampling theorem can cause the convergence curve to run abruptly, e.g., a “zig-zag” in the left region of the convergence curves Fig. 5-8, 10-13. This, too, is a characteristic of the behavior of the method at the limit of its applicability. In fact, calculated model (too roughly discretized) can deviate greatly from the originally intended original due to large “stairs”. These “stairs” suggest abrupt transitions between the material parameters - with possible numerically caused reflections. These can show strongly varying global and local errors, (Bronstein et al, 2005), with possible “outliers” of the values for neighboring values of the discretization fineness. Because, even with a sufficiently fine discretization, such abrupt transitions can cause a strong increase of the errors, compare to (Pregla, 2008). That would most likely also mean that the resulting GTL

is mathematically stiff (section 4.2.6). In this paper, abrupt transitions serve the purpose of testing the methods under the most difficult conditions possible. This is why "zig-zag" convergence curves result with a very coarse discretization. It is important, however, that with an ever finer discretization, the course of the convergence curve very soon becomes perfectly monotonic. This is also an indication of the stability of the given method according to the initial values in the sense of (Bronstein et al, 2005). This behavior and thus the stability, consistency, and convergence of all methods have been confirmed again and again as a result of numerous tests with different structures.

4.2.5 The behavior in the region of a fine discretization

The investigated methods theoretically have different orders of convergence p and should also converge at different ratios. In addition, the convergence rates are influenced differently by the local stiffness and thus by the corresponding local discretization errors. However, this effect will show fewer and fewer differences between the methods with ever finer discretization. Finally, all methods converge to the exact solution, and the error converges to zero, which was also confirmed in the numerical experiment.

However, it should be noted that the convergence curves show different values of the global discretization error $O(\Delta\overline{u}^p)$ at the end of the common calculation window when plotted together in one diagram, e.g., Fig. 5, 8, 10 and 13. It should also be noted that the exact convergence value can only be estimated. For this purpose, a control calculation will be carried out with a significantly larger number of discretization points than can be found in the presented calculation window. The calculated value is assumed to be approximately the "convergence value".

In the practical implementation of one-step methods, a rounding error $O(1/\Delta\overline{u})$ can be added to the global discretization error $O(\Delta\overline{u}^p)$. As a result, the step length $\Delta\overline{u}$ should not be too small, (Bronstein et al, 2005), (Zeidler, 2004). However, in numerous tests of various structures, such an accumulation of rounding errors $O(1/\Delta\overline{u})$ was only observed in a single case, with certain, deliberately unfavorable conditions. In this case, a convergence was initially observed with a discretization that was not yet too fine. A further refinement of the discretization first showed a slow convergence, which eventually turned into a slow divergence. This fact confirms that the rate of convergence is affected by errors. Despite an apparently low probability of occurrence, the user should take this possibility into account and not choose the length of the step too small. A practical check is known to be difficult to implement; see corresponding numerical tests in (Bronstein et al, 2005) or (Zeidler, 2004). A possible cumulation of a rounding error $O(1/\Delta\overline{u})$ can noticeably slow down the rate of convergence in the region of finer discretization.

4.2.6 Mathematical stiffness

The mathematical stiffness should not necessarily be the characteristic of a coarse discretization, as already described in section 4.2.4. There is currently no generally applicable definition of stiffness. In general, differential equations are mathematically "stiff" if they contain some constructs or parameters that cause rapid variations in the solutions. It is generally difficult to integrate "stiff" equations by ordinary numerical methods. Small errors may rapidly accumulate. See, e.g., (Curtiss and Hirschfelder, 1952).

With regard to the MoL, the definition of the stiffness in (Bronstein et al, 2005) appears to be suitable: An ODE is stiff if its solutions are made up of different, strongly exponentially decreasing components. In other words, a stiffness occurs when there is a large difference in scale on the same task. It may well happen that certain strongly decreasing components hardly make a contribution to the solution, but have a significant influence on the choice of the step size $\Delta\overline{u}$, so that the egg flow of the rounding error $O(1/\Delta\overline{u})$ increases very strongly, (Bronstein et al, 2005). In this case, the equation can cause a particularly high computational effort or, in extreme cases, especially with an adaptive choice of the step size, force the user to abort the calculation because of the apparent "standstill" of the calculation. In the sense, the stiff differential equations can certainly pose challenges with regard to the success of the solutions. Much more important in the sense of the paper, however, are the following obvious facts:

1. An ODE stiffness is generally difficult to identify, mostly only when there is a noticeably strong impact. A problem can be recognized as stiff if, for example, a convergence is noticeably slow and a computational time is long (e.g. a discretization error is large).
2. An ODE can show the stiffness to a variable and difficult to identify extent: The equation can be "more or less stiff" and accordingly more or less brings a noticeable regular error into the solution and, - depending on the degree of stiffness, - more or less extends the calculation time in order to achieve a given accuracy. For a suggestion on how the accuracy can be determined by the user, see section 4.4.5. If an implicit method

is used, the calculation time also depends - and to a considerable extent - on the stiffness, but also on the specific control parameters of the iterations. The result is that the computation time - according to the degree of stiffness - has less informative value about a specific method but rather reflects the conditions of the specific solution (i.e., a specific distribution of the material parameters of the structure being examined): Due to a local stiffness, the computation time can fluctuate strongly in an almost uncontrollable manner depending on the local material parameters, see also section 4.4.3.

3. In the case of a waveguide: The more modes participate in the overall field distribution (or in the overall impedance distribution), the more likely the stiffness of the GTL is. Specific cases of abrupt changes in the field distribution are difficult to predict. They can arise as a superposition of several modes and depend on many of their parameters, e.g., amplitudes, phases and spatial field distributions of the individual modes.
4. The influence of the stiffness of an ODE can depend, among other things, on the values of certain parameters, e.g., on the step size $\Delta\overline{u}$, e.g., see section 4.4.3. A too coarse discretization can cause a local stiffness and thus large fluctuations in the computation time due to the resulting abrupt transitions in material parameters ("stairs").

4.2.7 Some waveguide specifics

As discussed in Section 4.1, complex structures can exhibit complex spatial (and temporal) field distributions that cannot be easily predicted, even when only a few modes are involved.

Another specific aspect occurs relatively rarely but can occasionally falsify the final result of the calculation if handled improperly. This is a change in the composition of the modes caused by certain factors during repeated calculations of the eigenvalue problem. The cause can be an already very small change of certain input parameters, e.g., a slightly different (or differently placed) longitudinal discretization. This can result in an incorrect calculation of the final result because the wrong mode is used - or the previously used mode is no longer capable of propagation under the new conditions. The background of the effect from the point of view of the numerics is the following. The number of eigenvalues computed as a solution to an eigenvalue problem is usually equal to the number of discretization points along the cross-section of the structure. The eigenvalues correspond to the eigenmodes of the waveguide. But only some of them are capable of propagation and are not evanescent (if the real part of the propagation constant is zero). If the user wants to excite one or more modes capable of propagation (section 4.3), he should select these modes in the matrices appropriately and correctly distinguish them from the others eigenmodes in subsequent calculations. This can be done in two ways, either by the sequence number of the eigenvalues or by the shape of the corresponding field distribution along the cross-section of the structure (the values of this field distribution are contained in the corresponding column of the eigenvector matrix that has the same sequence number). If the calculations of the eigenvalue problem are repeated with slightly changed input parameters (e.g., with a different number of longitudinal discretization points), it can happen that the eigenmode previously assumed to be guiding now has a different sequence number - or even becomes unable to propagate. The consequences of such a "sudden" change, a reallocation, which is usually unexpected for the user - let's call it a "mode jump" - can be fatal for the correctness of the calculation: The software routine, which is usually responsible for recognizing the guiding modes, leads now carries out further calculations with an incorrect eigenmode - or no longer finds a suitable eigenmode at all. This phenomenon also occurs in the Floquet domain: A reallocation of the Floquet modes or a loss of their ability to propagate (the criterion of propagation ability is $Re\{\boldsymbol{G}_{\mathrm{F}}\} = 0$, where $\boldsymbol{G}_{\mathrm{F}}$ is the Floquet modes phase, (Pregla, 2008)).

It is reasonable to assume that the effect of the "mode jumps" (among other things) is closely related to the mathematical stiffness of solutions, e.g., with a coarser discretization, abrupt transitions of the material parameters and their high contrasts.

4.3 Test procedure

To verify the function of the methods as part of the MoL-IAFT-FD and to recognize some of their properties, some test structures are used. Dielectric, periodic and infinitely long 2D test structures are used for simplicity and efficiency. The structures are analyzed by calculating Floquet modes for only one period (Helfert and Pregla, 1998), (Pregla, 2004), (Pregla, 2008). However, the corresponding parameters of the eigenmodes of a period, i.e., Floquet modes, as well as transformation rules, must first be determined. They then apply to the beginning of A and the end of B of each period.

For the purpose of further simplification, a symmetry of the material parameters in the direction of propagation is considered. Otherwise, the spatial and the longitudinal distribution of the material parameters is selected

empirically so that the quality of the interpolation can be tested on it, e.g., abrupt and smooth transitions in the spatial distribution of the material parameters.

The testing procedure is as follows. A period of a structure is 2D discretized - in the propagation direction and across it. The test consists of several identical test calculations. With each next calculation, the number of discretization points in the propagation direction is increased, i.e., $N_i \in [20,30,40,...,1000]$. For each i-th calculation, a reference excitation $\widetilde{\mathbf{E}}_{\mathrm{A,f}}$ that is the same for all calculations is assumed and the resulting energy transported through the periodic structure $S^{(\mathrm{P})}_{\mathrm{A,B}}$ is examined. If the FD calculation is stable, the values of the energy transport $S^{(\mathrm{P})}_{\mathrm{A,B}} = f(N_i)$ converge to a value that is characteristic of the given structure. The computational time to achieve a given rate of convergence or a fixed error and/or the course of the convergence can serve as a quality features for each interpolation method used. The subscript "f" stand for a forward propagation part of the entire field.

The Floquet-modal matrices can be calculated from the eigenvalue problem

$$\boldsymbol{L}_{\mathrm{h}} = \boldsymbol{S}_{\mathrm{E}}^{-1} \boldsymbol{z}_{11\mathrm{h}} \boldsymbol{y}_{11\mathrm{h}} \boldsymbol{S}_{\mathrm{E}} = \boldsymbol{S}_{\mathrm{H}}^{-1} \boldsymbol{y}_{11\mathrm{h}} \boldsymbol{z}_{11\mathrm{h}} \boldsymbol{S}_{\mathrm{H}} \tag{73}$$

very efficiently by using the open- and short circuit matrix parameters of half the periods (Pregla, 2008) (s. "Impedance/admittance transformation" in section 2.3). One determine the Floquet modes from the input impedance and admittance without the need of a matrix inversion. Therefore, these expressions are numerically stable, (Pregla, 2008). The subscript "h" symbolizes half of the period. In the case of the symmetrical period, $z_{22} = z_{11}$ and $z_{12} = z_{21}$. The relation between $\boldsymbol{L}_{\mathrm{h}}$ and the Floquet modes phase $\boldsymbol{G}_{\mathrm{F}}$ is:

$$\tanh\left(\frac{1}{2}\boldsymbol{G}_{\mathrm{F}}\right) = \boldsymbol{L}_{\mathrm{h}}^{-\frac{1}{2}} = \boldsymbol{I} / \sqrt{\boldsymbol{S}_{\mathrm{E}}^{-1} \boldsymbol{z}_{11\mathrm{h}} \boldsymbol{y}_{11\mathrm{h}} \boldsymbol{S}_{\mathrm{E}}} = \boldsymbol{I} / \sqrt{\boldsymbol{S}_{\mathrm{H}}^{-1} \boldsymbol{y}_{11\mathrm{h}} \boldsymbol{z}_{11\mathrm{h}} \boldsymbol{S}_{\mathrm{H}}} \tag{74}$$

The characteristic wave impedance of any structure in the Floquet domain is an adequate unit matrix $\widetilde{\boldsymbol{Z}}_0 = \boldsymbol{I}$. It is transformed into the original domain using the Floquet-modal matrices $\boldsymbol{S}_{\mathrm{E}}$, $\boldsymbol{S}_{\mathrm{H}}$ specific to each structure:

$$\boldsymbol{Z}_0 = \boldsymbol{S}_{\mathrm{E}} \boldsymbol{I} \boldsymbol{S}_{\mathrm{H}}^{-1} \tag{75}$$

The excitation is simulated by the vector of the forward propagating Floquet modes:

$$\widetilde{\mathbf{E}}_{\mathrm{A,f}} = [1,0,\cdots,0]^t \tag{76}$$

The fundamental Floquet mode is used as a test field. An arrangement of the Floquet-modal matrix $\boldsymbol{S}_{\mathrm{E,H}}$ is assumed that the fundamental mode can be found in the first column. The vector $\widetilde{\mathbf{E}}_{\mathrm{A,f}}$ is also transformed into the original domain:

$$\mathbf{E}_{\mathrm{A,f}} = \boldsymbol{S}_{\mathrm{E}} \widetilde{\mathbf{E}}_{\mathrm{A,f}} \tag{77}$$

Before that, the z-parameters of the analyzed structure are calculated using the impedance/admittance transformation. The following applies to the input impedance:

$$\boldsymbol{Z}_{\mathrm{A}} = z_{11} - z_{12} \left(z_{22} + \boldsymbol{Z}_0\right)^{-1} z_{21} \tag{78}$$

The magnetic field distribution at input A is

$$\mathbf{H}_{\mathrm{A}} = 2 \left(\boldsymbol{Z}_{\mathrm{A}} + \boldsymbol{Z}_0\right)^{-1} \mathbf{E}_{\mathrm{A,f}} \quad \mathbf{E}_{\mathrm{A}} = \boldsymbol{Z}_{\mathrm{A}} \mathbf{H}_{\mathrm{A}} \tag{79}$$

and the energy flux density (energy transport) from the Poynting vector

$$S^{(\mathrm{P})}_{\mathrm{A}} = \mathbf{E}^t_{\mathrm{A}} \mathbf{H}^*_{\mathrm{A}} \tag{80}$$

The calculation for input A is sufficient because the output should ideally have exactly the same values. However, the additional calculation (the same values) for output B can be useful for verification purposes. The field distribution at output B would then be

$$\mathbf{H}_{\mathrm{B}} = \left(z_{11} + \boldsymbol{Z}_0\right)^{-1} z_{12} \mathbf{H}_{\mathrm{A}} \quad \mathbf{E}_{\mathrm{B}} = \boldsymbol{Z}_0 \mathbf{H}_{\mathrm{B}} \tag{81}$$

and the energy transport at the output of the periodic section is

$$S^{(\mathrm{P})}_{\mathrm{B}} = \mathbf{E}^t_{\mathrm{B}} \mathbf{H}^*_{\mathrm{B}} \tag{82}$$

4.4 Numerical results

4.4.1 Test conditions

In the paper, periodic structures were used for waveguiding tests. So it seemed appropriate to use the test excitation in the form of a specific Floquet mode. If it were not the Floquet domain but the eigenmode domain, it would be appropriate to choose the fundamental mode, for example. The peculiarity of the Floquet modes in comparison to eigenmodes consists in an indefiniteness as to which Floquet mode may be described as "fundamental". If for an eigenmode (in the so-called mode domain according to (Pregla, 2008)), the size of the real or imaginer part of the propagation constant can serve as a unique characteristic, a Floquet mode corresponds to several eigenmodes in the mode domain. Their propagation constants cannot be determined unambiguously because the propagation constant of the Floquet mode already contains the length of the Floquet period and the corresponding phase may have already exceeded the full circulation of 2π. So, if only one Floquet mode is used for a test, one can decide relatively freely which Floquet mode can be assumed to be "fundamental". In the paper, a specific Floquet mode was selected for each test structure based on the shape of its field distribution:

In the case of a defect waveguide in a photonic crystal, it was the "guiding" Floquet mode, which is characterized by a dominant main lobe in cross section of the structure corresponding to the defect waveguide, see Fig. 9, right. The parameters of the photonic crystal corresponded to the bandgap for the working wavelength.

In the case of the simple structures, a Floquet mode was assumed to be "guiding" if it had its field distribution in the form of half a sinusoid (see Fig. 2 and 4, each on the right).

All three test structures (Fig. 2, 4 and 9, left and center) have a width of 16.108 μm and a length of the Floquet period of 3.1 μm. These dimensions were specifically chosen to enable the largest possible number of eigenmodes and thus not to distort the Floquet modes used. As expected, the enlargement of the structure width in the numerical experiment made a possibly finer discretization necessary - because of more and more propagating modes with finer spatial field distribution. The width was discretized to 200 points for all structures. The number of discretization points in the longitudinal direction was varied to determine the convergence curves. The refractive index varies in the range from 1.0 to 3.4. The wavelength $\lambda = 1550$ nm. The test procedure has already been described in section 4.3.

The convergence curves are expected to have different values of the global discretization error $O(\Delta\overline{u}^p)$ at the end of the common calculation window when plotted together in one diagram. The convergence of the error to zero cannot be shown in more detail due to the limitation of the linearly scaled calculation window. However, the use of logarithmic scaling did not seem appropriate in the paper because the two regions of discretization refinement are to be shown in detail using only one diagram: the coarser and the finer discretization. It should also be noted that the exact convergence value can only be estimated. For this purpose, a control calculation will be carried out with a much larger number of discretization points than the presented calculation window can have (usually 5 times larger). The calculated value is assumed to be approximately the "convergence value".

4.4.2 Results: General findings

All tested methods show convergence, thus are stable and therefore demonstrably consistent. With a finer discretization, the numerical solutions for all investigated methods converge to the exact solution. Finally, the error converges to zero.

All methods of interpolation, both the two already established and all-new, deliver on the whole comparable results. However, depending on the method, it turned out that the computing time to reach a fixed error differed greatly. This enables an effective comparison, which, however, is only valid for specific structural parameters and for the specific method. For the purpose of comparison, some examples of the different computational times for some concrete structures are presented. Besides, one should consider that the computational time for the implicit method GRK4 also depends on the end criteria of the respective iterations at each step.

The results, in particular the course of the convergence, are generally qualitatively comparable to those in (Pregla, 2006-b) or (Pregla, 2008). However, in the example in (Pregla, 2006-b), there should be significantly fewer modes capable of propagation.

It was also found that the convergence rates can obviously be influenced differently by the local stiffness and thus by corresponding local discretization errors.

Accordingly, the following effects have been observed for different waveguiding structures in numerical experiments:

- All methods were tested under the same numerically difficult conditions. In general, all methods, old and new, show a similar reaction to the respective difficulties, e.g., abrupt transitions of the material parameters.

This can serve as an indication of the correct implementation of the individual methods. Each difference is presented in more detail below.

– Single “outliers” of the final results, especially in the region of the coarse discretization, i.e., relatively strong differences in the final result for individual discretization densities compared to the smaller and larger neighboring values. These “outliers” are able to “wander” along the axis of the discretization density, depending on the variation of other specified parameters. If the corresponding guiding mode remains propagable (the real part of the Floquet propagation constant is zero), the cause of these “outliers” appears to be local discretization density dependent mathematical stiffness, with a large, resulting global error.
  A good example is the “zig-zag” shape of the convergence curves in all test structures with abrupt transitions of the material parameters and their relatively high contrast: the convergence curves in Fig. 5-8 (the test structure with a symmetrical step profile of the permittivity, Fig. 4) and Fig. 10-13 (the defect waveguide in photonic crystal, Fig. 9). An example of an unfavorable superposition of the individual eigenmodes is the single “outlier” of the convergence curve in Fig. 12 (at the number of discretization points 550). This shows how a small change in the discretization density has caused a visibly different superposition of the (only slightly changed) eigenmodes. However, it is important to note that the irregular shape of all convergence curves generally becomes smoother as the discretization density increases.
  It should be noted that the test structure with the smooth profile of the permittivity does not have any “zig-zag” shapes (Fig. 3), because there is obviously also a smooth field distribution, and thus, no significant stiffness is present.
  The test structure with a rectangular shape (Fig. 4) shows a higher irregularity of convergence curves than the test structure with a photonic crystal (Fig. 9), with the same contrast of abrupt transitions of material parameters. This can obviously be explained by the fact that the wave guidance mainly takes place through the defect waveguide. The wave propagation through the photonic crystal is significantly prevented by the bandgap at the working wavelength. As a result, the influence of the discretized elements of the photonic crystal on the resulting wave propagation is much smaller.

4.4.3 Results for some test structures

As an example, the following three test structures (Fig. 2, 4, 9) and their corresponding convergence curves (Fig. 3, 5-8, 10-13) are shown below.

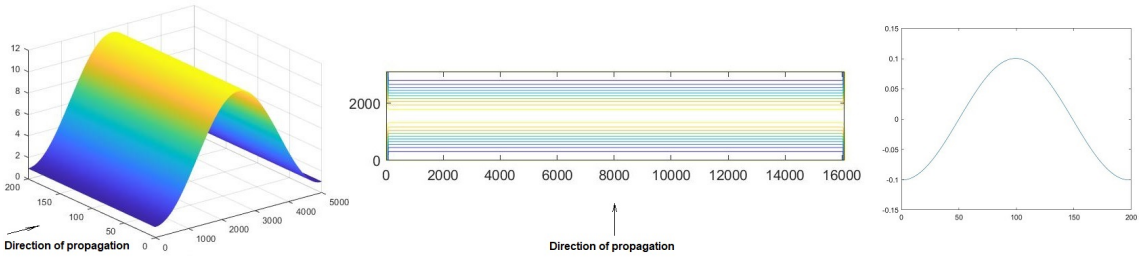


Fig. 2 2D test structure with a cosine profile of the permittivity distribution in lateral direction. Only one period is shown. The arrows show the direction of the wave propagation. The Neumann boundary conditions for the normal component of the electric field were assumed on the left and right sides of the structure. Left: The vertical axis shows the permittivity $\varepsilon_r = n^2$, the refractive index $n = 1.0 \div 3.4$. The horizontal axes represent the sampling points of the spatial discretization. The scale of the longitudinal axis is shown stretched. Center: The view from above in the original scale. The two axes represent the dimensions of the structure in [nm]. The propagation occurs in the vertical direction. Right: Lateral field distribution of the used Floquet mode in the cross-section of the structure (a.u.).

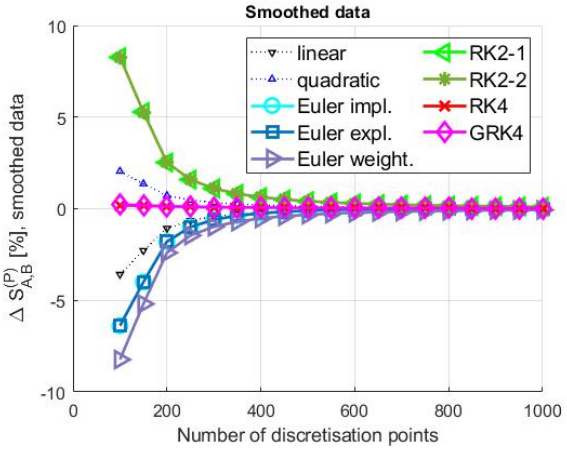


Fig. 3 All convergence curves of the "cosine" test structure in Fig. 2. This structure is characterized by rather smooth and steady convergence curves. All convergence curves are so smooth that occasional additional smoothing (shown on this diagram) does not reveal any significant differences from the original curves. The efficiency of the methods differs. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK2-1 and RK2-2 roughly coincide. The same applies to Euler-impl and Euler-expl. All curves converge to zero in the further course.

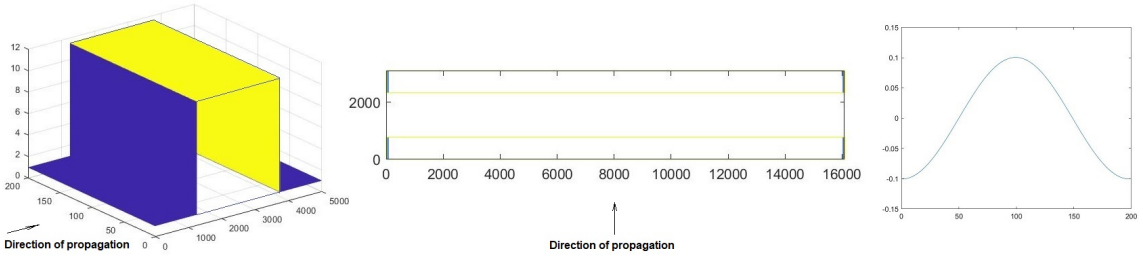


Fig. 4 2D test structure with a symmetrical step profile of the permittivity distribution in lateral direction. Only one period is shown. The arrows show the direction of wave propagation. The Neumann boundary conditions for the normal component of the electric field were assumed on the left and right sides. Left: The vertical axis shows the permittivity $\varepsilon_r = n^2$, the refractive index $n = 1.0 \div 3.4$. The horizontal axes represent the sampling points of the spatial discretization. The scale of the longitudinal axis is shown stretched. Center: The view from above in the original scale. The two axes represent the dimensions of the structure in [nm]. The propagation occurs in the vertical direction. Right: Lateral field distribution of the used Floquet mode in the cross-section of the structure (a.u.).

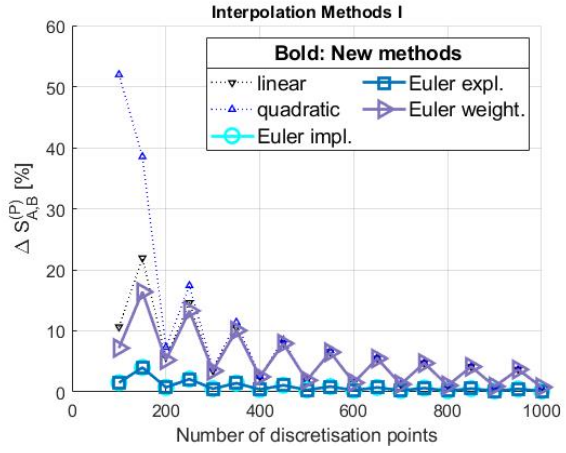


Fig. 5 Five convergence curves of the "pulse" test structure in Fig. 4. This structure is characterized by rather zig-zag convergence curves. The zig-zag curve shows an obvious dependence of the result on small changes in the discretization grid. A complex interplay between the spatial periodicities of the eigenmodes and the placement of the discretization points can be assumed. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of Euler-impl and Euler-expl roughly coincide. All curves converge to zero in the further course.

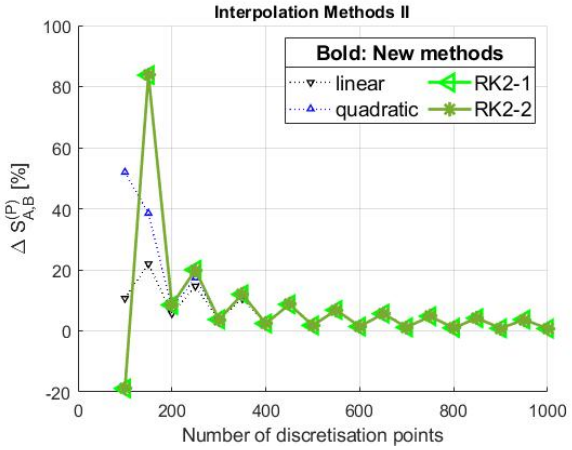


Fig. 6 Further convergence curves of the test structure in Fig. 4 ("pulse"). The convergence curves for conventional linear and quadratic interpolation are shown for comparison. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK2-1 and RK2-2 roughly coincide. All curves converge to zero in the further course.

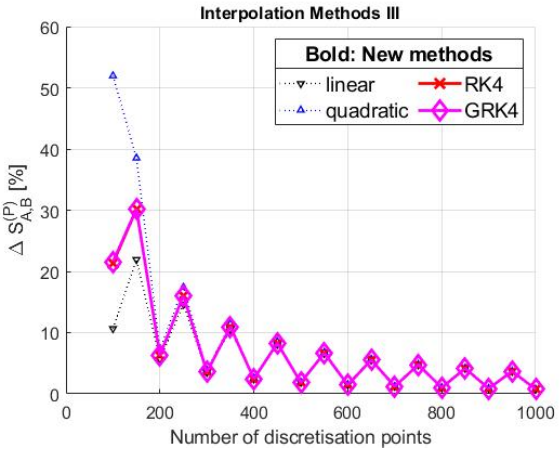


Fig. 7 Further convergence curves of the test structure in Fig. 4 ("pulse"). The convergence curves for conventional linear and quadratic interpolation are shown for comparison. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK4 and GRK4 roughly coincide. All curves converge to zero in the further course.

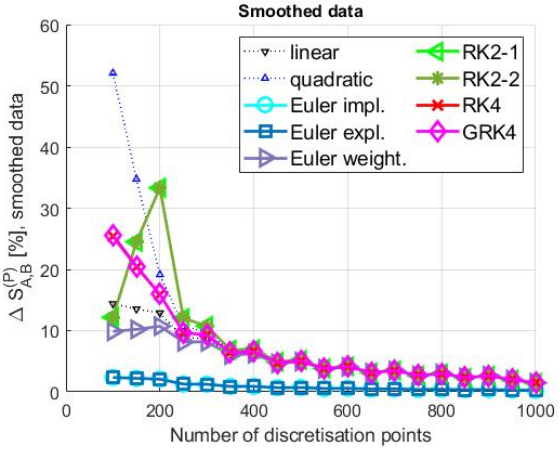


Fig. 8 All convergence curves of the test structure in Fig. 4 ("pulse"). Smoothing has been applied for better visibility of the actual convergence. The smoothing is carried out using the Savitzky-Golay filter method by the range of 20% of the total number of data points. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK2-1 and RK2-2 roughly coincide. The same applies to Euler-impl and Euler-expl as well as to RK4 and GRK4. All original curves converge to zero in the further course.

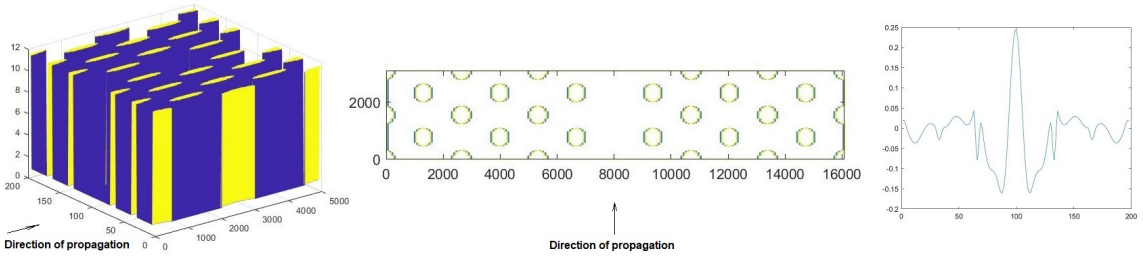


Fig. 9 2D test structure as a defect waveguide in a photonic crystal with round dielectric rods ($n = 3.4$) in air. Only one period is shown. The arrows show the direction of wave propagation. The dielectric rods with the diameter $2r = 0.4a$, where $a = 1550$ nm is the crystal constant. The Neumann boundary conditions for the normal component of the electric field were assumed on the left and right sides. Left: The vertical axis shows the permittivity $\varepsilon_r = n^2$. The horizontal axes represent the sampling points of the spatial discretization. The scale of the longitudinal axis is shown stretched. Center: The view from above in the original scale. The two axes represent the dimensions of the structure in [nm]. The propagation occurs in the vertical direction. Right: Lateral field distribution of the used Floquet mode in the cross-section of the structure (a.u.).

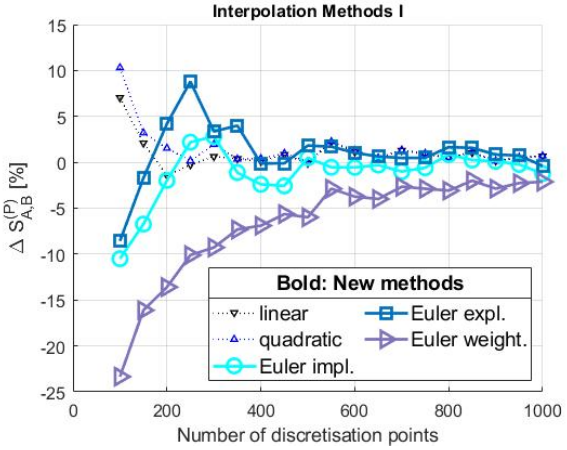


Fig. 10 Some convergence curves of the test structure in Fig. 9 (defect waveguide in a photonic crystal). The zig-zag curve is similar to the case of the symmetrical pulse in Fig. 5-7 but is obviously weakened because of an even more complex interference of the eigenmodes and/or numerical reflections (compare to (Pregla, 2008)). The wave guidance mainly takes place through the defect waveguide. The wave propagation through the photonic crystal is significantly prevented by the bandgap at the working wavelength. As a result, the influence of the discretized elements of the photonic crystal on the resulting wave propagation is much smaller. One sees the periodic nature of the error (there are three slightly pronounced lobes: in the center of the diagram, as well as on the left and right). The structure of the defect waveguide in a photonic crystal is much more complex than the pulse structure. In this test, the above-mentioned effects presumably occur in a significantly larger number. In this way, the effect of a statistical "smoothing effect" is possible. For comparison of computational time, see table 1 in section 4.4.4. All curves (including those for Euler-weight) converge to zero in the further course.

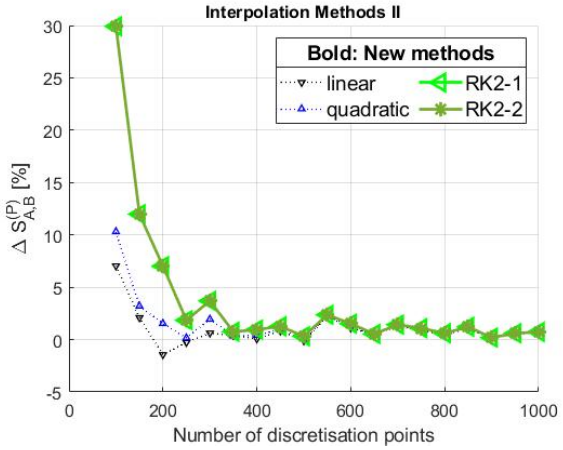


Fig. 11 Further convergence curves of the test structure in Fig. 9 (defect waveguide in a photonic crystal). An increase in the global error in the center of the diagram probably has to do with a less favorable superposition of the eigenmodes and/or numerical reflections. One sees the periodic nature of the error (there are three slightly pronounced lobes: in the center of the diagram, as well as on the left and right). The convergence curves for conventional linear and quadratic interpolation are shown for comparison. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK2-1 and RK2-2 roughly coincide. All curves converge to zero in the further course.

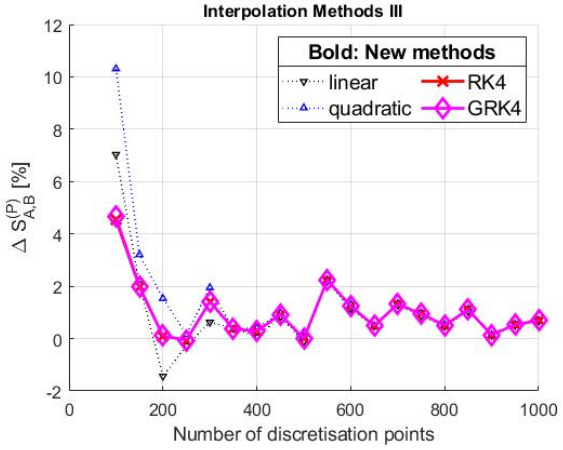


Fig. 12 Further convergence curves of the test structure in Fig. 9 (defect waveguide in a photonic crystal). An increase in the global error in the center of the diagram probably has to do with a less favorable superposition of the eigenmodes and/or numerical reflections. One sees the periodic nature of the error (there are three slightly pronounced lobes: in the center of the diagram, as well as on the left and right). The convergence curves for conventional linear and quadratic interpolation are shown for comparison. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK4 and GRK4 roughly coincide. All curves converge to zero in the further course.

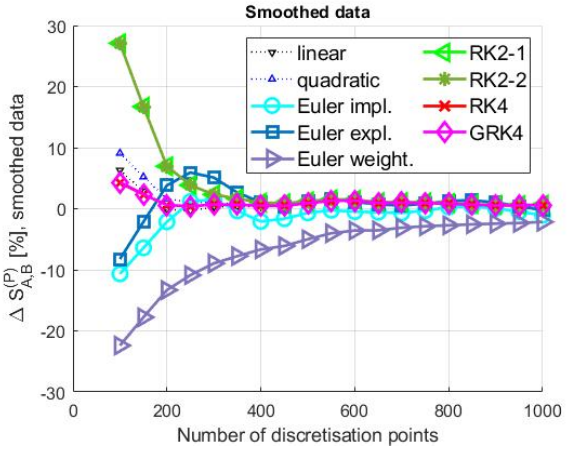


Fig. 13 All convergence curves of the test structure in Fig. 9 (defect waveguide in a photonic crystal). Smoothing has been applied for better visibility of the actual convergence. The smoothing is carried out using the Savitzky-Golay filter method by the range of 20% of the total number of data points. For comparison of computational time, see table 1 in section 4.4.4. The convergence curves of RK2-1 and RK2-2 roughly coincide. The same applies to RK4 and GRK4. All original curves converge to zero in the further course.

4.4.4 About the computational time and rate of convergence

The next aspect is the different suitability of different methods for different use cases. Depending on the method and the structure, the computational time to reach a fixed error was measured (see table).

Table 1 Computational time "t" to reach a fixed error vs the number of discretization points "N". Content of a cell: [N / t]. "<": The specified error convergence value was already reached at the smallest tested discretization density.

Cosine profile, $\Delta S_{A,B}^{(P)} = 1\ \%$

Linear	Euler-impl	Euler-expl	RK2-1	RK2-2	Euler-weight	RK4	Quad	GRK4
200 / 4.2 s	250 / 3.3 s	250 / 5.4 s	350 / 9.6 s	350 / 9.4 s	400 / 14.0 s	<100 / 2.1 s	250 / 9.3 s	<100 / 8.0 s

Pulse profile, $\Delta S_{A,B}^{(P)} = 10\ \%$

Linear	Euler-impl	Euler-expl	RK2-1	RK2-2	Euler-weight	RK4	Quad	GRK4
250 / 5.9 s	<100 / 1.3 s	<100 / 1.7 s	300 / 7.9 s	300 / 8.0 s	200 / 5.2 s	250 / 8.1 s	250 / 9.3 s	250 / 16.2 s

Photonic crystal profile, $\Delta S_{A,B}^{(P)} = 5\ \%$

Linear	Euler-impl	Euler-expl	RK2-1	RK2-2	Euler-weight	RK4	Quad	GRK4
150 / 3.1 s	150 / 1.8 s	250 / 5.4 s	250 / 6.5 s	250 / 6.4 s	500 / 17.9 s	100 / 2.1 s	150 / 3.6 s	100 / 8.5 s

After a detailed examination of the table and the corresponding convergence curves, it can be stated that only very general conclusions can be drawn from them (section 4.4.5). The other options do not appear to be practicable with reasonable effort. The reasons for this have already been discussed in section 4.1. Further research may be necessary.

4.4.5 Initial systematics and recommendations for users

One can assume that in the case of a complex multimode test structure, a combination of several disturbing effects can occur, e.g., physical and numerical reflections, stiffness of the solution, cumulative rounding errors, ill-placed discretization and unsuitable boundary conditions. In this way, the effects of the individual causes can mask each other. Therefore, it is difficult to derive the corresponding systematics about the overall performance of a certain method from a limited number of exploratory tests. Therefore, only very general and certainly incomplete systematics can be derived in the paper. Their completion appears to be a subject for further research.
As a result of the test analysis, three perspectives appear to make sense:

1. For a given test structure (profile): How fast were the methods? - The ranking of a computational time to reach a given error.
2. For a given method: What features were shown? Which applications are better suited?
3. General recommendations.

1. For a given test structure (profile): How fast were the methods?

Cosine profile: One can assume that the model represents smooth solutions in general.

RK4 (GRK4)*	Euler-impl	Linear	Euler-expl	RK2-2 ≈ Quad	RK2-1	Euler-weight

Notice: *GRK4: Fast convergence, but the computation time is influenced by the end criteria of the iterations. According to computation time, the conspicuously best methods are RK4, Euler-impl and Linear.

Pulse profile:

Euler-impl	Euler-expl	Euler-weight	Linear	RK2-1 ≈ RK2-2 ≈ RK4	Quad	GRK4

Notice: Linear was noticeably slower this time, by about 5 times.

Photonic crystal profile:

Euler-impl ≈ RK4	Linear	Quad	Euler-expl	RK2-2	RK2-1	GRK4	Euler-weight

Notice: RK4 was just as fast as Euler-impl and Linear in this case, while Euler-weight was about 8 times slower.

2. For a given method: What features were shown? Which applications are better suited?
** Sorted in ascending order according to the accumulated computation time for all three structures (as a score and as an estimated overall performance). See also the relevant remarks.*

1. *Euler-impl:* (6 s) The fastest and most universal method in the test. Possibly has great potential in possible applications that is yet to be investigated.
2. *RK4:* (12 s) A universal method, especially suitable for smooth and mixed solutions. Shows remarkably fast convergence for smooth solutions. Possibly has great potential in possible applications that is yet to be investigated.
3. *Euler-expl:* (12 s) Possibly universal. It can serve as an alternative to the methods mentioned above.
4. *Linear:* (13 s) Relatively universal, but may be inferior to Euler-imp for steep solutions in the test by about 5 times.
5. *Quad:* (22 s) Relatively universal and relatively fast, it was inferior to the Euler-imp by about 3.5 times in overall performance.
6. *RK2-2:* (23 s) Appears to be similar to Quad.
7. *RK2-1:* (25 s) Appears to be similar to Quad and RK2-2.
8. *GRK4:* (33 s) The method has a special status: Its computation time becomes decisive through the choice of the final accuracy of the method-internal interpolations at each step. With a generally long computation time, the method has special properties, e.g. more tolerant to stiffness, which should be further investigated in application to waveguides. Shows remarkably fast convergence for smooth solutions. Recommended for smooth solutions, but especially for unclear problematic cases.
9. *Euler-weight:* (37 s) Particularly suitable for steep solutions. The method also has a special status: Its computation time and other properties can be changed continuously from the properties of the Euler-expl to the properties of the Euler-impl by choosing the weighting. As a result, the method appears to have great potential in possible applications.

3. These other recommendations for the user appear to be useful:

- If linear distributions of the material parameters with abrupt transitions dominate, the user can first apply the methods of the 1st order: Euler-impl, Euler-expl or Linear. The user should expect possible negative effects of the stiffness: Certain software routines should automatically intercept such events, issue adequate informative messages and adapt the further course of the calculations to the users' requirements.
- If the distributions of the material parameters are mostly smooth, it would be expected that the higher-order methods would provide the best results with minimal computation time: Especially the fast RK4. If problems arise due to stiffness, GRK4 is recommended.
- If there is no information about the expected solutions, an efficient exploratory approach is recommended (see above): For this purpose, the user can first create a library of different interpolation methods as encapsulated software routines. There should be a possibility of easy, flexible switching between the individual methods, i.e., between the corresponding software routines. Individual control routines can measure the calculation time of the individual steps, classify them and make them available for viewing as required. The library of the individual methods should be connected to "the outside world", i.e., the remaining mathematical software components, through a uniform interface. The interface should be able to interact with all required mathematical software.
- In the case of numerical problems with a specific structure, the procedure of numerical experiment - or in other words numerical testing - can be based on already sufficiently accumulated experience, including in other branches. An example is the practice of software testing, (ISTQB, 2021), (Spillner and Linz, 2019), which has well developed over the past decade. Accordingly, for example, an efficient exploratory approach with elements of the variational analysis can be helpful: In a specific case, the user should carry out the initial calculations of his task using a certain method, observe the intermediate results and evaluate them with regard to his requirements and the possible influencing factors. From this follows the direction of the further procedure: As an example, the changes in the end result can be examined as a function of small changes in the discretization density or a different spatial placement of the discretization points. Such an approach (or type of test) is often the only alternative for complex systems with diverse system states, e.g., in the software testing, if no helpful systematic of the test object can be derived beforehand, compare to (ISTQB, 2021) and (Spillner and Linz, 2019).

The additional recommendations in case the user has to work in the region of coarser discretization:

– Numerical experiments show that for any given structure, each method has its individual critical discretization density or sampling rate. Beyond this limit, the method remains "stable" but no longer delivers authentic results due to the heavily distorted model. The user does not have to determine the actual limit of the sampling theorem separately. This limit can easily be recognized by the relatively strong variations of the final results for closely adjacent values of the discretization density. A few test runs of the calculation are sufficient for this. If the calculations are conspicuously complex, one may be able to avoid full calculation by examining intermediate results in an exploratory systematic way.

The additional recommendations for the region of finer discretization:

– A cumulation of the rounding error $O(1/\Delta\overline{u})$ has so far been extremely rare in our numerical experiments. The corresponding effects are relatively difficult to recognize (see corresponding numerical tests in (Bronstein et al, 2005) or (Zeidler, 2004)), which is also associated with a large numerical effort. A cumulation of a rounding error $O(1/\Delta\overline{u})$ can noticeably slow down the rate of convergence in the region of finer discretization. Despite an apparently low probability of occurrence, the user should take this possibility into account and not choose the length of the step too small.

5 Conclusions

An enhancement of instruments for solution of general transmission line equations with method of lines, impedance/admittance and field transformation in combination with finite differences has been proposed. Several methods have been integrated into the framework of the IAFT-MoL-FD. This enables the potential user to analyze the widest possible type of waveguide structures. The newly integrated methods only serve as an example: In this way, other one-step and multistep methods from numerical mathematics can also be integrated. This option increases the usability of the MoL-IAFT-FD for potential users.

All investigated methods ensure that with a finer discretization the numerical solution converges to the exact solution. All methods of interpolation, both the two already established and all-new, deliver on the whole comparable results.

Some insights were given into the issues of stability, convergence and consistency, including a clarification of the relevant definitions that seem appropriate for the engineering of the MoL applications.

The behavior of the corresponding solutions in the regions of a coarse and a finer discretization was discussed, especially for the purpose of analyzing the wave-guiding structures.

Some insights were given into the question of mathematical stiffness, which was hardly mentioned before in the context of the MoL-IAFT-FD, including a discussion of the possible impact on the efficiency of the analysis of the waveguiding structures. Some effects in the context of waveguides that have been little mentioned so far but are well known in the practice were discussed, e.g., the "mode jumps". It has been confirmed that the rate of convergence is affected by the stiffness and thus by local discretization errors and rounding errors $O(1/\Delta\overline{u})$.

Depending on the method, it turned out that the computing time to reach a fixed error - and thus the efficiency of a concrete method in a concrete application - differed greatly. This enables an effective comparison. However, this efficiency can change considerably for a different concrete structure or its scaling. Although the range of all possible applications requires an individual, application-specific choice of a suitable method, general recommendations for the user have been derived. This selection can be made with the help of an experience-based exploratory test.

The results are generally qualitatively comparable to those in the literature.

Therefore, a broader choice of different methods is a helpful aid when there are increased demands on the efficiency, i.e., the computation time or on the stability of the solution.

Acknowledgment

The analysis presented here was based on the excellent theory and earlier works by Univ. Prof. Dr. Reinhold Pregla. His advice is greatly appreciated. Also, I thank Dr. Stefan F. Helfert for his valuable advice and for sharing his rich experience.

References

Barcz A, Helfert S F and Pregla R (2002) Modelling of 2D photonic crystalls by using the method of lines. In: Europ. Symp. on Photonic Crystals (ESPC 2002), (21.-25.04.2002, Warsaw, Poland), 2002.

Bronstein I N, Semendjaew K A, Musiol G and Mühlig H (2005) Taschenbuch der Mathematik, Verlag Harri Deutsch, 2005.

Bultheel A and Cools R (2010) The Birth of Numerical Analysis, World Scientific, 2010.

Chen W K (2004) The electrical engineering handbook, Elsevier, 2004.

Conradi O Helfert S F and Pregla R (2001) Comprehensive Modeling of Vertical–Cavity Laser–Diodes by the Method of Lines, IEEE J. Quantum Electron., vol. 37, pp. 928–935, 2001.

Curtiss C F and Hirschfelder J O. (1952) Integration of Stiff Equations. Proc. Nat. Acad. Sci. of the U S. A. 38:235–243, 1952.

Dahlquist G (1963) A special stability problem for linear multistep methods, BIT 3, 27-43, 1963.

Greda L (2004) Neue Diskretisierungsschemata zur genauen Analyse komplexer dreidimensionaler Strukturen der Mikrowellentechnik und Optik. PhD Thesis, VDI Verlag, Fortschritt-Berichte VDI, Reihe 21, nr. 369, Hagen, 2015.

Grothmann R (2015) Skriptum Numerik. *https://docplayer.org/13570032-Skriptum-numerik-prof-dr-rene-grothmann.html*, 2015.

Grothmann R (2012) Skriptum Numerik. *http://www.sozialinformatik.de/fileadmin/_migrated/content_uploads/Numerik-Skript.pdf*, 2012.

Hairer E, Nørsett S P and Wanner G (1993) Solving ordinary differential equations. 1, Nonstiff problems. Springer Verlag, 1993.

Hairer E and Wanner G (1996) Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, Springer 2, 1996.

Hamdi S, Schiesser W E, Griffiths G W (2007) Method of lines. Scholarpedia, 2(7):2859, doi:10.4249/scholarpedia.2859, 2007.

Helfert S F, Barcz A and Pregla R (2003) Three-dimensional vectorial analysis of waveguide structures with the method of lines. Springer, Optical and quantum electronics, vol. 35, nr. 4, pp. 381–394, 2003.

Helfert S F A and Pregla R (1998) Efficient analysis of periodic structures. J. Lightwave Technol., vol. 16, no. 9, pp. 1694–1702, Sep. 1998.

Helfert S F A and Pregla R (2002) The method of lines: a versatile tool for the analysis of waveguide structures. Electromagnetics, vol. 22, pp. 615–637, 2002. Invited paper for the special issue on "Optical wave propagation in guiding structures".

Helfert S F A and Pregla R (1996) Finite Difference Expressions for Arbitrarily Positioned Dielectric Steps in Waveguide Structures, J. Lightwave Technol., vol. 14, no. 10, pp. 2414–2421, Oct. 1996.

Hildebrand F B (1987) Introduction to Numerical Analysis. 2nd ed., Dover, New York, 1987.

Hoellig K (2011) Numerische Methoden fuer Differenzialgleichungen, Universität Stuttgart, Fachbereich Mathematik (Ed.), *www.mathematik-online.org*, 2011.

Hoellig K (1998) Grundlagen der Numerik, MathText, 1998.

International Software Testing Qualifications Board (2021) https://www.istqb.org/downloads.html, 2021.

Pregla R and Helfert S F (2002) Modeling of Microwave devices with the method of lines. Recent Research developments in Microwave Theory & Techniques, pp. 145–196, Research Signpost, Kerala, India, 2002.

Pregla R (2002) Modeling of optical Waveguide Structures with general anisotropy in arbitrary orthogonal coordinate systems. IEEE Journal of selected topics in quantum electronics, vol. 8, nr. 6, pp. 1217–1224, 2002.

Pregla R (2004) Analysis of gratings with symmetrical and unsymmetrical periods. IEEE, Proceedings of 2004 6th International Conference on Transparent Optical Networks (IEEE Cat. No. 04EX804), vol. 1, pp. 101–104, 2004.

Pregla R (2006) Analysis of Microvawe Structures by Combination of the Method of Lines and Finite Differences. In: MICON (08.-09.04.2006, Crasow, Poland), 2006.

Pregla R (2006) Modeling of optical waveguides and devices by combination of the method of lines and finite differences of second order accuracy. In: Optical and Quantum Electronics, 38 pp. 3–17, 2006.

Pregla R and Pascher W (1989) The Method of Lines, in Numerical Techniques for Microvawe and Millimeter Wave Passive Structures. T. Itoh, (Ed.), pp. 381–446, J. Wiley Publ., New York, USA, 1989.

Pregla R (1999) Efficient and Accurate Modeling of Planar Microwave Structures with Anisotropic Layers by the Method of Lines (MoL). in Int. Symp. on Recent Advances in Microwave Technology, Malaga, Spain, pp. 699–703, Dec. 1999.

Pregla R (2002) Efficient and Accurate Modeling of Planar Anisotropic Microwave Structures by the Method of Lines. IEEE Trans. Microwave Theory Tech., vol. 50, no. 6, pp. 1469-1479, June 2002.

Pregla R (2006) Modeling of optical waveguides and devices by combination of the method of lines and finite differences of second order accuracy. Optical and Quantum Electronics 38:3–17, 2006.

Pregla R (2003) Efficient Modeling of Periodic Structures. AEU, vol. 57, pp. 185–189, 2003.

Pregla R (2008) Analysis of Electromagnetic Fields and Waves - The Method of Lines. Wiley & Sons, Chichester, UK, 2008.

Pulch R (2020) Numerik Grundpraktikum: Numerische Verfahren für gewöhnliche Differentialgleichungen, *https://math-inf.uni-greifswald.de/storages/uni-greifswald/fakultaet/mnf/mathinf/pulch/skript_numerik_gp.pdf*, 2020.

Samarski A A (1982) Introduction to the numerical Methods (in russ.), Nauka, 1982.

Samarski A A (1986) Introducción a los Métodos Numéricos, Mir, 1986.

Schiesser W E (1991) The Numerical Method of Lines Integration of Partial Differential Equations, Academic Press, San Diego , USA, 1991.

Schiesser W E and Griffiths G W (2009) A compendium of Partial Differential Equation Models: Method of Lines Analysis with Matlab. Cambridge University Press, 2009.

Seyrich J K H (2016) Numerical Integrators for Physical Applications. PhD thesis, Universität Tübingen, 2016.

Spiller W, Helfert S F and Jahns J (2019) The numerical investigation of colliding optical solitons as an all-optical-gate using the method of Lines. Springer, Optical and Quantum Electronics, vol. 51, nr. 5, pp. 1–22, 2019.

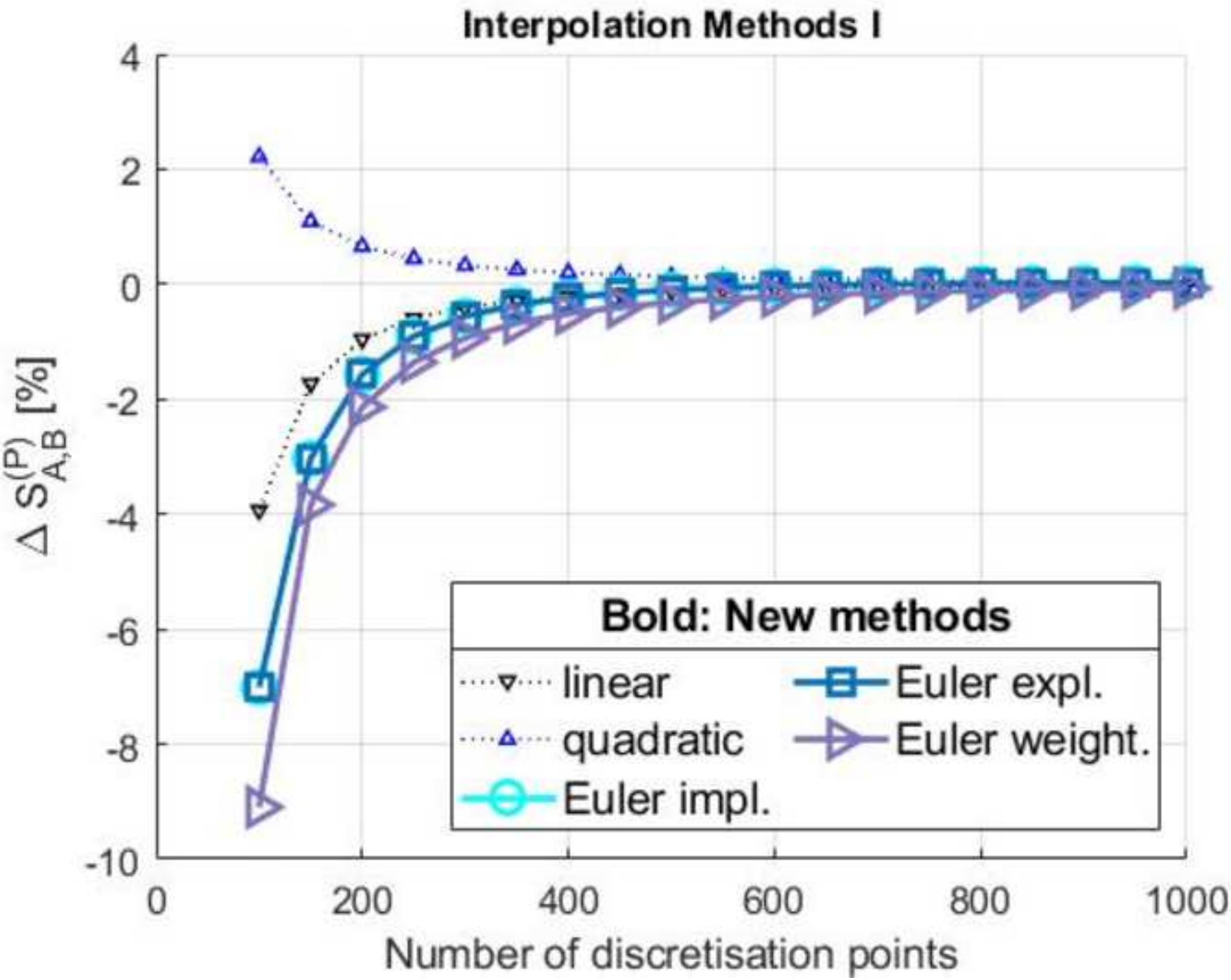
Spiller W (2022) Optimization of the analysis of waveguides bends and Y-junctions with the method of lines, to be published.

Spillner A and Linz T (2019) Basiswissen Softwaretest: Aus- und Weiterbildung zum Certified Tester - Foundation Level nach ISTQB ® - Standard, dpunkt. verlag, (2019).

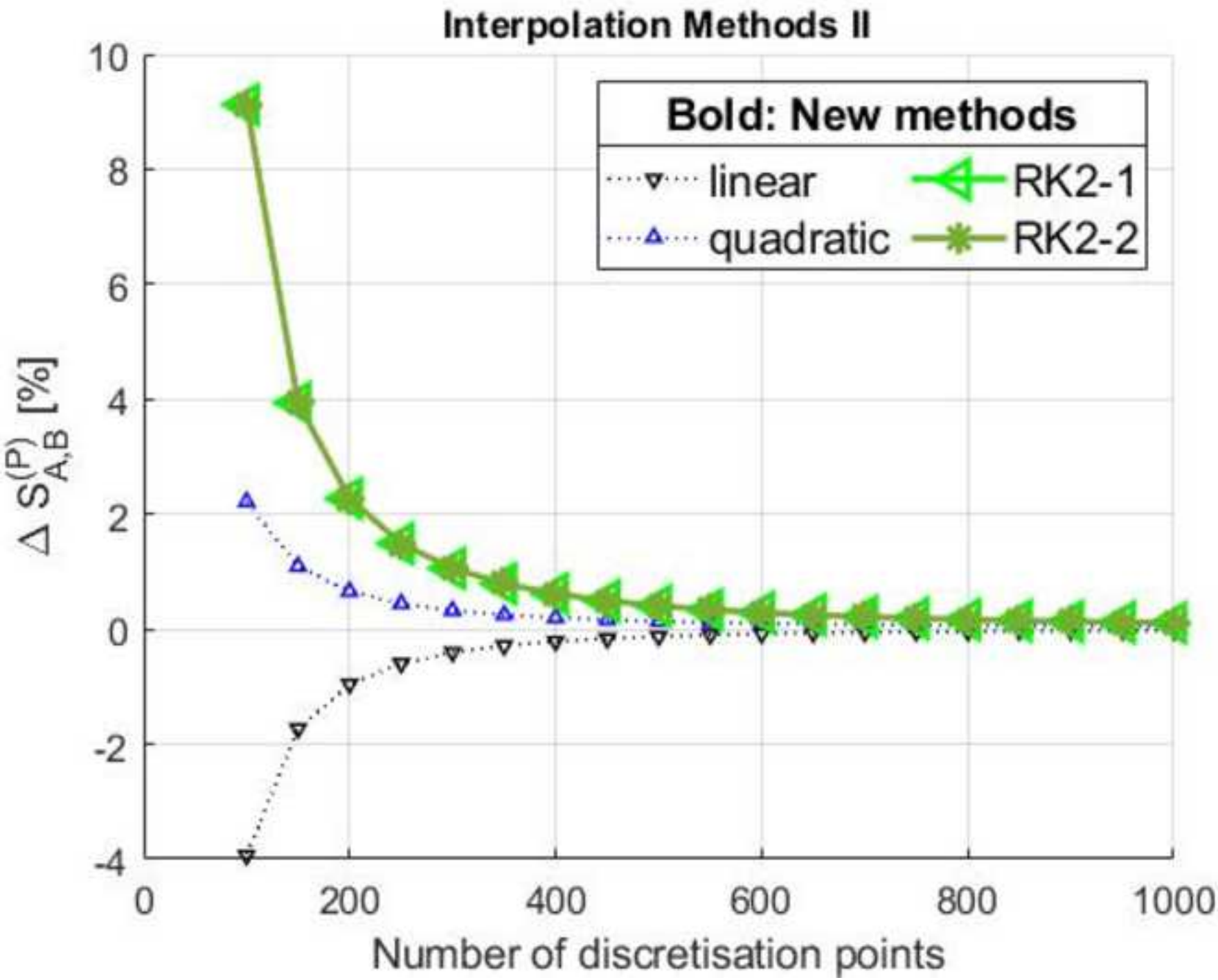
Vietzorreck L (2001) Numerical Simulation of Three-Dimensional MMIC Discontinuities by the Method of Lines. PhD thesis, FernUniversität Hagen, 2001.

Zeidler E (2004) Oxford Users' Guide to Mathematics. Oxford University Press, 2004.

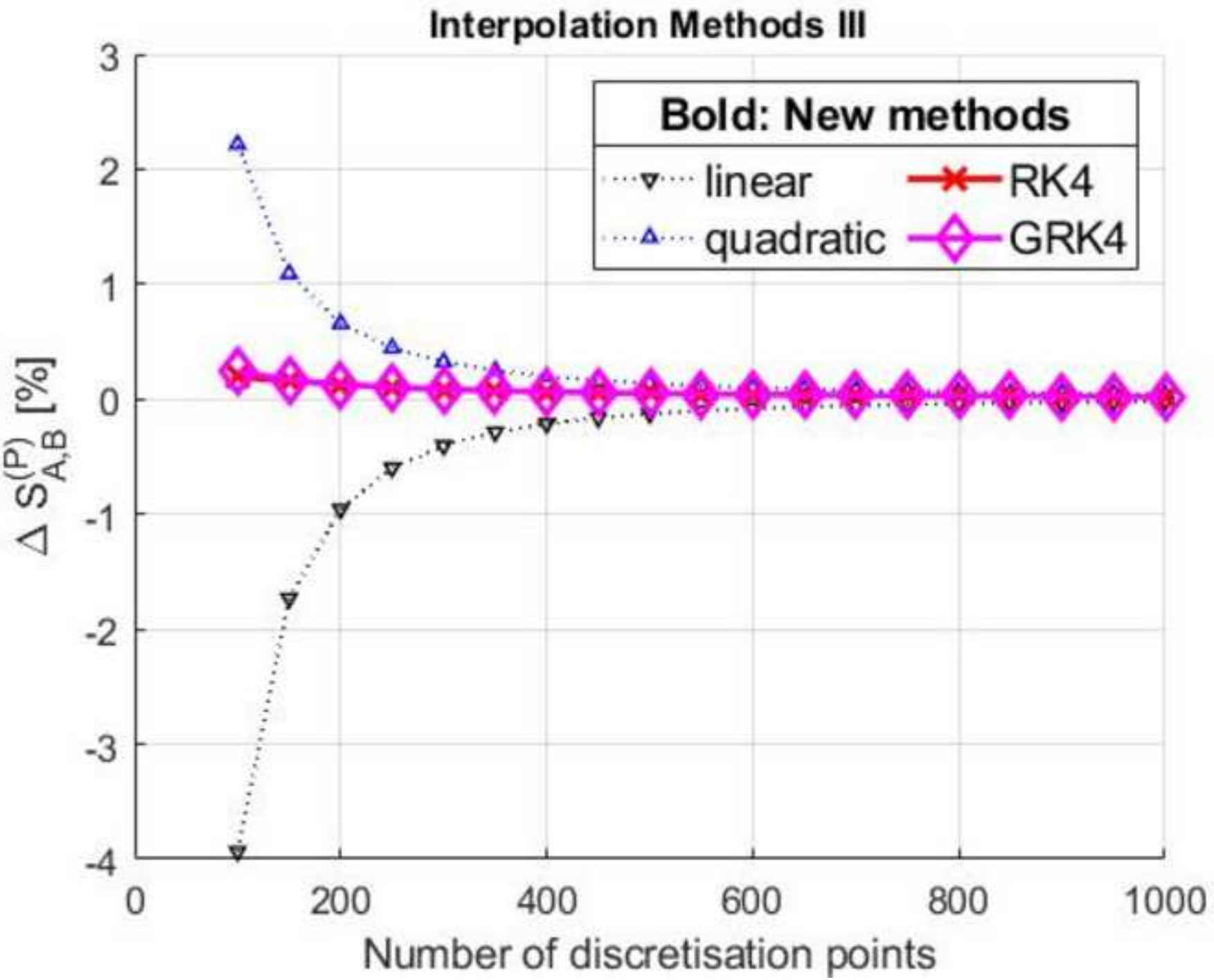
Figure


Interpolation Methods I
4
2
0
-2
-4
-6
-8
-10
0
200
400
600
800
1000
Number of discretisation points
Δ S(P) A,B [%]
Bold: New methods
linear
quadratic
Euler impl.
Euler expl.
Euler weight.

Figure


Interpolation Methods II
Bold: New methods
linear
quadratic
RK2-1
RK2-2
$\Delta S^{(P)}_{A,B}$ [%]
10
8
6
4
2
0
-2
-4
0
200
400
600
800
1000
Number of discretisation points

Figure


Interpolation Methods III
Bold: New methods
linear
quadratic
RK4
GRK4
$\Delta S^{(P)}_{A,B}$ [%]
3
2
1
0
-1
-2
-3
-4
0
200
400
600
800
1000
Number of discretisation points

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Example.eps