Institut für Angewandte und
Numerische Mathematik

Technische Universität Wien

Report Nr. 77/88

Asymptotic Error Expansions for Stiff Equations: The Implicit Midpoint Rule

W. Auzinger
R. Frank

Asymptotic Error Expansions for Stiff Equations: The Implicit Midpoint Rule

W. Auzinger and R. Frank *

Abstract

This is a study of the structure of the global discretization error of the implicit midpoint rule applied to nonlinear stiff initial value problems. Those special cases where an asymptotically correct expansion in even powers of the stepsize h exists have been described in an earlier paper. In the general stiff case the global error does not admit a "pure" asymptotic expansion but there occur dominant oscillating terms at the h^2-level. The amplitude of these oscillations shows a very smooth behavior and, consequently, the error structure is sufficiently regular to guarantee a satisfactory efficiency of acceleration algorithms like extrapolation or defect correction. These algorithmic applications are discussed in a separate paper.

1 Introduction

Asymptotic expansions of the global discretization error in powers of the stepsize h have been well known for a long time for many types of problems and discretization methods (cf. for instance [8], [14]). If, for instance, the implicit midpoint rule

$$\begin{aligned} \frac{1}{h}\Big(\zeta_\nu - \zeta_{\nu-1}\Big) &= f\Big(\frac{1}{2}(t_{\nu-1}+t_\nu), \frac{1}{2}(\zeta_{\nu-1}+\zeta_\nu)\Big), \\ \zeta_0 &= z_0 \end{aligned} \tag{1.1}$$

or the implicit trapezoidal rule

$$\begin{aligned} \frac{1}{h}\Big(\zeta_\nu - \zeta_{\nu-1}\Big) &= \frac{1}{2}\Big(f(t_{\nu-1},\zeta_{\nu-1}) + f(t_\nu,\zeta_\nu)\Big), \\ \zeta_0 &= z_0 \end{aligned} \tag{1.2}$$

are applied to an initial value problem of ordinary differential equations,

$$\begin{aligned} y'(t) &= f(t,y(t)), \\ y(0) &= z_0 \end{aligned} \tag{1.3}$$

with the exact solution $z(t)$, then - under suitable smoothness assumptions w.r.t. (1.3) - the global discretization error at $t_\nu = \nu h$ admits an "even" expansion

$$\zeta_\nu - z(t_\nu) = h^2 e_2(t_\nu) + h^4 e_4(t_\nu) + \ldots + h^{2q} e_{2q}(t_\nu) + R_\nu \tag{1.4}$$

*Institut fuer Angewandte und Numerische Mathematik, Technische Universitaet Wien, Wiedner Hauptstrasse 6–10, A-1040 Wien, Austria.

where the functions $e_{2i}(t)$ are solutions of certain linearized initial value problems - the so-called "variational equations". The expansion (1.4) is asymptotically correct, i.e. there holds

$$\|R_\nu\| \le E(t_\nu; L, \ldots)\, h^{2q+2} \tag{1.5}$$

where $E(\ldots)$ depends (smoothly) on certain parameters which characterize the given problem (1.3), namely on the Lipschitz bound L of the right hand side f of (1.3), on certain bounds for higher derivatives of f, etc. For nonstiff problems with smooth data functions these problem-characterizing parameters are of moderate size and therefore (1.4) and (1.5) describe the structure of the global error in a satisfactory way. The implicit methods (1.1) and (1.2), however, are designed to solve (nonsmooth) *stiff* initial value problems with large Lipschitz constants $L \gg 0$, for which the quantity $E(t; L, \ldots)$ in (1.5) is very large; therefore the estimate (1.5) is worthless in the stiff case.

Recently, asymptotic error expansions for stiff problems have been investigated by several authors (cf. [1],[2],[9],[15]). In [2], in particular, stiff situations have been characterized for which the expansion (1.4) for the global error of the midpoint and trapezoidal rules is asymptotically correct, i.e. where

$$\|R_\nu\| \le E \cdot h^{2q+2} \tag{1.6}$$

is valid with a moderate quantity E depending not on the Lipschitz constant L but only on those parameters that remain moderate for the stiff problem under consideration. The assumptions for the proof of (1.6) were:

(i) $\varepsilon \le Ch^{2q}$ ("strongly stiff case"), where C denotes some moderate constant and where $\varepsilon > 0$ is a small parameter characterizing the stiffness (i.e., the order of magnitude of the stiff eigenvalues is $-1/\varepsilon$).

(ii) It has to be assumed that - except f_y which is of course $O(\varepsilon^{-1})$ - a certain number of derivatives of the right hand side f are moderate, i.e., $O(\varepsilon^0)$.

(iii) Certain *coupling conditions* between the stiff and the non-stiff components are required.

Concerning these assumptions, the trapezoidal rule has turned out to be more "robust" than the midpoint rule: The smoothness requirements (ii) w.r.t. f are milder in the case of the trapezoidal rule; moreover the short, asymptotically correct expansion

$$\zeta_\nu - z(t_\nu) = h^2 e_2(t_\nu) + O(h^4) \quad \text{for} \quad \varepsilon \le Ch^2 \tag{1.7}$$

exists - in contrast to the midpoint rule - independent of any coupling condition.

The present paper is devoted to a discussion of global error structures in situations where not all assumptions (i)–(iii) hold and where therefore the validity of (1.6) cannot be expected. Since for the trapezoidal rule, (1.6) can be concluded for a larger class of problems, a further study of the midpoint rule, which is "less robust" w.r.t. the validity of (1.6), appears to be of particular interest. In the present paper we therefore restrict our considerations to the midpoint rule; it should, however, be noted that a similar type of analysis can also be carried out for the trapezoidal rule or other symmetric schemes like the semi-implicit midpoint rule proposed in [5].

The following simple example illustrates the situation:

Example. Consider the scalar model problem of Prothero and Robinson,

$$\begin{aligned} y'(t) &= -\frac{1}{\varepsilon}\Big(y(t) - g(t)\Big) + g'(t), \\ y(0) &= g(0) \quad (= z_0) \end{aligned} \tag{1.8}$$

with a smooth function $g(t)$. Obviously, $z(t) \equiv g(t)$ is the solution of (1.8). Due to the "nonsmooth" inhomogeneity $\frac{1}{\varepsilon}g(t) + g'(t)$, the right hand side of (1.8) does not satisfy the above smoothness assumptions (ii) required for asymptotic correctness in the case of the midpoint rule. Indeed, it is well known (cf. for instance Example 2 in [2]) that the global error $\zeta_\nu - z(t_\nu)$ of the implicit midpoint rule applied to (1.8) involves *nonsmooth terms at the h^2-level* which contain the factor

$$\left(\frac{1-\frac{h}{2\varepsilon}}{1+\frac{h}{2\varepsilon}}\right)^{\nu} \quad (\approx (-1)^\nu \text{ for } \varepsilon \ll h). \tag{1.9}$$

The occurrence of such terms at the h^2-level inhibits the existence of an asymptotically correct expansion. The behavior of the nonsmooth terms (1.9) depends on the ratio between ε and h:

- $\varepsilon \le Ch^2$: Here the nonsmooth terms show a very regular, *oscillating* behavior: (1.9) can be written as
$$\left(\frac{1-\frac{h}{2\varepsilon}}{1+\frac{h}{2\varepsilon}}\right)^{\nu} = \left(\frac{1-\frac{2\varepsilon}{h}}{1+\frac{2\varepsilon}{h}}\right)^{\nu} \cdot (-1)^\nu \tag{1.10}$$
and for fixed $t = t_\nu = \nu h$ the amplitude of this oscillation reads
$$\left(\frac{1-\frac{2\varepsilon}{h}}{1+\frac{2\varepsilon}{h}}\right)^{t/h} = \exp\left(\frac{t}{h}\ln\left(\frac{1-\frac{2\varepsilon}{h}}{1+\frac{2\varepsilon}{h}}\right)\right) \approx \exp\left(-\frac{4\varepsilon}{h^2}t\right), \tag{1.11}$$
i.e., for $\varepsilon \le Ch^2$ this amplitude can be described by a smooth, h-dependent function.
- $Ch^2 \le \varepsilon \le Ch$: In this case the amplitude of (1.10) is no longer smooth but shows a significantly decaying behavior; with decreasing h this damping effect becomes stronger.
- $h \le C\varepsilon$: For $h \le 2\varepsilon$ the terms (1.9) are no longer oscillating but show a strongly decaying behavior. For fixed $t = t_\nu = \nu h$ and for $h \to 0$,
$$\left(\frac{1-\frac{h}{2\varepsilon}}{1+\frac{h}{2\varepsilon}}\right)^{\nu} = \exp\left(\frac{t}{h}\ln\left(\frac{1-\frac{h}{2\varepsilon}}{1+\frac{h}{2\varepsilon}}\right)\right) \approx \exp\left(-\frac{t}{\varepsilon}\right). \tag{1.12}$$

In the present paper we extend the results of [2] in a twofold way:

a) We describe the structure of the global error for a more general class of nonlinear stiff systems, where it will turn out that nonsmooth terms - similar as the terms (1.9) in the above example - occur[1] at the dominant h^2-level. We give a quantitative description of these nonsmooth terms.

b) We relax the requirement of "strong stiffness" $\varepsilon \le Ch^{2q}$; "mildly stiff" situations are also considered.

According to the behavior of the typical terms (1.9) for different constellations of ε and h, the "ε-h-plane" divides in a natural way into several regions. In sections 3 and 4 we shall discuss the

[1] In those special cases which have been discussed in [2] these terms are also present but with an additional factor ε. This is exactly the reason why, for those cases, the desired estimate (1.6) could be shown to be valid in the strongly stiff case $\varepsilon \le Ch^{2q}$. In the mildly stiff case, however (i.e., if $\varepsilon \le Ch^{2q}$ is not satisfied), this additional factor ε is not sufficient to ensure the asymptotic correctness of (1.4).

cases $\varepsilon \leq Ch^2$ and $\varepsilon \leq Ch$, respectively. The case $h \leq C\varepsilon$ will not be discussed here but can be treated with similar arguments as were given in [1], subsection 5.3, for the implicit Euler scheme.

It is an inevitable fact that "full" asymptotic expansions cannot be maintained in the general stiff case (as is already obvious from the above example). Our assertions about the structure of the global error of the midpoint rule are in the following sense weaker than (1.4), (1.6):

- The expansion is not asymptotically correct in the sense of (1.6): The "remainder term" R_ν is not $O(h^{2q+2})$ but contains nonsmooth terms at the $O(h^2)$-level. However, these terms show a very regular behavior: For $\varepsilon \leq Ch^2$ they are oscillating with a smoothly varying amplitude (cf. Theorem 3.1, (3.30)); for $Ch^2 \leq \varepsilon \leq Ch$ these oscillations are superposed by certain damping effects (cf. Theorem 4.1, (4.25)).

- The solutions $e_{2i}(t)$ of the variational equations, which describe the smooth terms of the global error, are not h-independent but depend on h^2 in the following way: In our representation of the global error for $\varepsilon \leq Ch^2$ (cf. section 3) the $e_{2i}(t)$ depend smoothly on the moderate parameter ε/h^2; in our representation for $\varepsilon \leq Ch$ (cf. section 4) they depend smoothly on $(\varepsilon/h)^2$. The main reason for this dependence is the following: To describe the error structure for nonequidistant grids (which is of course essential in the context of stiff problems) we subdivide the integration interval into subintervals with constant stepsizes h and apply our results - which are formulated for equidistant grids (cf. Theorems 3.1 and 4.1) - in an inductive way. For each of these subintervals the starting values for the $e_{2i}(t)$ are influenced by the global error accumulated up to the end of the preceding subinterval. Theorem 3.1 or 4.1 - applied to the preceding subinterval - shows that the accumulated global error at the begin of the current subinterval depends smoothly on ε/h^2 or $(\varepsilon/h)^2$, resp.[2] So the starting values for the e_{2i} and therefore the e_{2i} themselves depend on these quantities.

Some remarks concerning the usefulness of our results are in order. The crucial question is whether the error structure is sufficiently regular to ensure an efficient performance of acceleration algorithms like extrapolation or defect correction – or whether the h^2-oscillations impair the efficiency of these methods. In a separate paper [3] we discuss extrapolation and defect correction for stiff problems, the results of the present paper forming the basis for these investigations. It turns out that - due to the regular behavior of the oscillating error components - acceleration algorithms are highly efficient and indeed belong to the class of very promising stiff solvers. Concerning extrapolation, for instance, it is shown that the h^2-oscillations with smooth amplitudes (for $\varepsilon \leq Ch^2$) are of no negative influence.[3]

Moreover, our results provide new insight concerning the question of suitable control mechanisms (control of stepsize and order, smoothing, ...): In mildly stiff situations, for instance, the higher extrapolation steps show a reduced order (at least immediately after the start or after a change of stepsize), and a stepsize and order control strategy should take account of this fact. Furthermore, our results enable the design of appropriate connection strategies. Global connection, for instance, seems to be favourable in the mildly stiff case (cf. [3] for a numerical illustration). An extensive analysis of these questions is under preparation.

[2]For the scalar example above, for instance, the dependence of the global error on ε/h^2 for $\varepsilon \leq Ch^2$ can be seen from (1.11): The amplitude of the oscillation at the h^2-level behaves like $\exp(-\frac{4\varepsilon}{h^2}t)$. Note that this dependence becomes insignificant with increasing stiffness, since $\exp(-\frac{4\varepsilon}{h^2}t)$ tends to the h-independent value 1 for $\varepsilon \to 0$.

[3]So far, extrapolation has only been analyzed in situations where pure asymptotic expansions in powers of h exist; in these cases the error is essentially polynomial in h, and so it is very easy to understand the effect of extrapolation. But for a satisfactory performance of extrapolation it is by no means necessary that the error is polynomial in h; any error behavior sufficiently smooth in h is equally appropriate.

2 Smooth Solutions of the Variational Equations. The Remainder Equation

Consider a stiff initial value problem (1.3) with the smooth solution $z(t)$. We assume that the stiffness of the given problem is characterized by a small parameter $\varepsilon > 0$, i.e.

$$f_y(t, z(t)) = T(t)\,\Lambda(t)\,T^{-1}(t), \qquad \Lambda(t) = \begin{pmatrix} c_1(t) & 0 \\ 0 & -\dfrac{c_2(t)}{\varepsilon} \end{pmatrix} \tag{2.1}$$

with smooth functions $T(t)$, $T^{-1}(t)$, $c_1(t)$ and $c_2(t)$. Let

$$Re\,(c_2(t)) \geq \kappa > 0 \tag{2.2}$$

where $1/\kappa$ is of moderate size independently of ε. We are considering the 2-dimensional case, i.e., $c_1(t)$ and $c_2(t)$ are scalar functions; but this is not a crucial restriction. All the material could easily be rewritten for problems which are more general in the sense that $c_1(t)$ and $c_2(t)$ are matrix functions of certain dimensions.[4]

The higher derivatives of $f(t, y)$ w.r.t. y are assumed to be smooth, i.e.

$$\begin{aligned} f_{yy}(t, y) &= O(\varepsilon^0), \\ f_{yyy}(t, y) &= O(\varepsilon^0), \\ &\vdots \end{aligned} \tag{2.3}$$

is assumed to be valid in a suitably defined region $\mathcal{R}$ containing the true solution $z(t)$ of (1.3) and its numerical (midpoint) approximation ζ_ν. The solvability of the algebraic equations (1.1), i.e. the existence of the ζ_ν within $\mathcal{R}$ is not discussed in the present paper; for a discussion of this point, cf. for instance [7].[5]

According to the ideas of Gragg [8] the global error of the implicit midpoint discretization (1.1) of (1.3) can be written as

$$\zeta_\nu - z(t_\nu) = h^2 e_2(t_\nu) + h^4 e_4(t_\nu) + \ldots + h^{2q} e_{2q}(t_\nu) + R_\nu \tag{2.4}$$

where the functions $e_2(t)$, $e_4(t)$, ... are solutions of the variational equations

$$\begin{aligned} e_2'(t) &= f_y(t, z(t))\, e_2(t) - \frac{1}{2^2 3!} z'''(t) + \frac{1}{2^2 2!} f_y(t, z(t))\, z''(t), \\ e_4'(t) &= f_y(t, z(t))\, e_4(t) - \frac{1}{2^4 5!} z^V(t) + \frac{1}{2^4 4!} f_y(t, z(t))\, z^{IV}(t) - \frac{1}{2^2 3!} e_2'''(t) + \\ &\quad + \frac{1}{2^2 2!} f_y(t, z(t))\, e_2''(t) + \frac{1}{2} f_{yy}(t, z(t)) \left(\frac{1}{2^2 2!} z''(t) + e_2(t) \right)^2 \\ &\vdots \end{aligned} \tag{2.5}$$

[4]In this case one would say that - in view of assumption (2.1) - there are two 'clusters' of eigenvalues (namely the eigenvalues of $c_1(t)$ and $-c_2(t)/\varepsilon$) representing the non-stiff and the stiff part of the spectrum, respectively.

[5]The following point is worth mentioning here: In most papers dealing with the question of solvability, it is assumed that the right hand side of the nonlinear stiff problem has a moderate (or even non-positive) one-sided Lipschitz constant. However, as has been demonstrated in [4], this assumption is violated for many stiff problems. Therefore a certain gap remains in the respective theory; but a discussion of this question is out of the scope of the present paper. We throughout assume that the ζ_ν exist within a suitable bounded neighborhood $\mathcal{R}$ of $z(t)$.

and where the remainder term R_ν satisfies the difference equation

$$\begin{aligned}\frac{1}{h}\Big(R_\nu - R_{\nu-1}\Big) \;&=\; f_y\Big(\hat t_\nu, z(\hat t_\nu)\Big)\,\frac{1}{2}\Big(R_{\nu-1}+R_\nu\Big) + \\ &\quad + G_\nu\left(\frac{1}{2}\Big(R_{\nu-1}+R_\nu\Big)\right) + f_y\Big(\hat t_\nu, z(\hat t_\nu)\Big)a_\nu + b_\nu - c_\nu .\end{aligned} \tag{2.6}$$

Here we have introduced the denotation

$$\hat t_\nu := \frac{1}{2}\Big(t_{\nu-1}+t_\nu\Big) = t_\nu - \frac{h}{2}. \tag{2.7}$$

For each ν, $G_\nu(\cdot)$ is a certain nonlinear function (cf. (2.11) below). The inhomogeneous terms a_ν and c_ν are defined as

$$a_\nu := \frac{1}{2}\left(I_\nu^- + I_\nu^+\right), \qquad c_\nu := \frac{1}{h}\left(I_\nu^+ - I_\nu^-\right) \tag{2.8}$$

where the $I^\pm_{\nu,2k}$ in $I^\pm_\nu := I^\pm_{\nu,0} + h^2 I^\pm_{\nu,2} + \ldots + h^{2q} I^\pm_{\nu,2q}$ are the remainder terms of the Taylor expansions

$$\begin{aligned} z\Big(\hat t_\nu \pm \frac{h}{2}\Big) \;&=\; z(\hat t_\nu) \pm \frac{h}{2} z'(\hat t_\nu) + \frac{h^2}{8} z''(\hat t_\nu) \pm \ldots \pm \\ &\qquad \pm \frac{h^{2q+1}}{2^{2q+1}(2q+1)!} z^{(2q+1)}(\hat t_\nu) + I^\pm_{\nu,0}, \\ e_2\Big(\hat t_\nu \pm \tfrac{h}{2}\Big) \;&=\; e_2(\hat t_\nu) \pm \frac{h}{2} e_2'(\hat t_\nu) + \frac{h^2}{8} e_2''(\hat t_\nu) \pm \ldots \pm \\ &\qquad \pm \frac{h^{2q-1}}{2^{2q-1}(2q-1)!} e_2^{(2q-1)}(\hat t_\nu) + I^\pm_{\nu,2}, \\ &\;\;\vdots \end{aligned} \tag{2.9}$$

Furthermore, using the abbreviation

$$\begin{aligned} h^2 u(t) \;:=&\; \frac{h^2}{2^2 2!} z''(t) + \frac{h^4}{2^4 4!} z^{IV}(t) + \ldots + \frac{h^{2q}}{2^{2q}(2q)!} z^{(2q)}(t) + \\ &+ h^2 e_2(t) + \frac{h^2}{2^2 2!} e_2''(t) + \ldots + \frac{h^{2q}}{2^{2q-2}(2q-2)!} e_2^{(2q-2)}(t) + \\ &\qquad + h^4 e_4(t) + \ldots + \frac{h^{2q}}{2^{2q-4}(2q-4)!} e_4^{(2q-4)}(t) + \ldots + \\ &\qquad\qquad + h^{2q} e_{2q}(t), \end{aligned} \tag{2.10}$$

we define b_ν as the collection of all terms of the Taylor expansion of $f(\hat t_\nu, z(\hat t_\nu) + h^2 u(\hat t_\nu) + a_\nu)$ about $(\hat t_\nu, z(\hat t_\nu))$ which contain at least a factor h^{2q+2}. So, by definition, a_ν, b_ν and c_ν contain the common factor h^{2q+2}.

The nonlinear functions G_ν appearing in (2.6) read

$$\begin{aligned} G_\nu(R) \;\equiv&\; h^2 G_\nu^{(1)}\cdot R + G_\nu^{(2)}(R)\cdot R^2 \;:= \\ :=&\; \int_0^1 f_{yy}\Big(\hat t_\nu, z(\hat t_\nu) + \sigma(h^2 u(\hat t_\nu) + a_\nu)\Big) d\sigma \cdot \Big(h^2 u(\hat t_\nu) + a_\nu\Big)\cdot R \;+ \\ &+ \int_0^1 f_{yy}\Big(\hat t_\nu, z(\hat t_\nu) + h^2 u(\hat t_\nu) + a_\nu + \sigma R\Big)(1-\sigma) d\sigma \cdot R^2 \end{aligned} \tag{2.11}$$

where the $G_\nu^{(1)}$ are linear; the factor h^2 within $h^2 G_\nu^{(1)}$ originates from $h^2 u(\hat{t}_\nu) + a_\nu$ (recall that a_ν contains even the factor h^{2q+2}).

All this can be shown by inserting $\zeta_\nu = z(t_\nu)+h^2 e_2(t_\nu)+\ldots+h^{2q}e_{2q}(t_\nu)+R_\nu$ (cf. (2.4)) into (1.2), using Taylor expansions about $\hat{t}_\nu = t_\nu - \frac{h}{2}$ and equating coefficients of powers of h. In the classical, nonstiff case, where all occurring derivatives are moderate, the inhomogeneity $f_y(\hat{t}_\nu, z(\hat{t}_\nu))a_\nu + b_\nu - c_\nu$ of the 'remainder equation' (2.6) is obviously $O(h^{2q+2})$ *with a moderate O-constant* and the asymptotic correctness of the expansion (2.4) follows immediately by a conventional stability estimate. In the stiff case, the variational equations are also stiff and therefore - if no special care is taken w.r.t. the starting values $e_{2i}(0)$ - the $e_{2i}(t)$ are not smooth (i.e., their derivatives are influenced by negative powers of ε). Following an idea of Dahlquist and Lindberg [6] we therefore base our structural analysis of the global error on *smooth solutions of the variational equations.* I.e., we shall construct special starting values $e_{2i}(0)$ such that the $e_{2i}(t)$ are sufficiently smooth to ensure that all derivatives appearing in a_ν, b_ν and c_ν are $O(1)$ and therefore a_ν, b_ν and c_ν in (2.6) are again $O(h^{2q+2})$. (Note that b_ν is not influenced by f_y and can therefore indeed be considered $O(h^{2q+2})$ for sufficiently smooth $e_{2i}(t)$.)

The crucial point is that the starting values $e_{2i}(0)$ for smooth solutions $e_{2i}(t)$ do usually not vanish. Since (2.4) has to be satisfied for $\nu = 0$, i.e.

$$0 = \zeta_0 - z(0) = h^2 e_2(0) + \ldots + h^{2q} e_{2q}(0) + R_0, \tag{2.12}$$

the starting value $R_0 = -h^2 e_2(0) - \ldots - h^{2q} e_{2q}(0)$ can therefore not be chosen at the desired $O(h^{2q+2})$-level but is only $O(h^2)$ in general. So the asymptotic correctness of (2.4) cannot be expected - and does indeed not hold in general. The main goal of the present investigation is the quantitative description of the $O(h^2)$-remainder term. (Only under special assumptions about the given problem the starting values $e_{2i}(0)$ contain an additional factor ε such that, for $\varepsilon \ll h$, $R_0 = O(h^{2q+2})$ is satisfied. These special cases have been studied in [2].)

Characterization of smooth solutions.

Our discussion of smooth solutions of the variational equations is based on singular perturbation techniques (similarly as in [1] and [2]).

We consider the first variational equation in (2.5) and transform it according to

$$\bar{e}_2(t) := T^{-1}(t) e_2(t) \tag{2.13}$$

(with $T(t)$ from (2.1)). This leads us to

$$\bar{e}_2'(t) \;=\; \Lambda(t)\bar{e}_2(t) + D(t)\bar{e}_2(t) + v_2(t) + \Lambda(t) w_2(t) \tag{2.14}$$

with the smooth inhomogeneities $v_2(t) := -\frac{1}{24}T^{-1}(t)z'''(t)$ and $w_2(t) := \frac{1}{8}T^{-1}(t)z''(t)$ and where

$$D(t) := -T^{-1}(t)T'(t) = \begin{pmatrix} d_{11}(t) & d_{12}(t) \\ d_{21}(t) & d_{22}(t) \end{pmatrix}. \tag{2.15}$$

We denote

$$\bar{e}_2(t) = \begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix}. \tag{2.16}$$

Due to the ε-structure of $\Lambda(t)$ (cf. (2.1)) the inhomogeneity of (2.14) is of the form

$$v_2(t) + \Lambda(t) w_2(t) = \begin{pmatrix} R_{0,2}(t) \\ \dfrac{1}{\varepsilon} S_{0,2}(t) + S_{1,2}(t) \end{pmatrix} \tag{2.17}$$

with certain smooth functions $R_{0,2}(t), S_{0,2}(t)$ and $S_{1,2}(t)$. Multiplying the second component of (2.14) by ε we obtain

$$\begin{aligned} x_2'(t) &= c_1(t)x_2(t) + d_{11}(t)x_2(t) + d_{12}(t)y_2(t) + R_{0,2}(t), \\ \varepsilon y_2'(t) &= -c_2(t)y_2(t) + \varepsilon d_{21}(t)x_2(t) + \varepsilon d_{22}(t)y_2(t) + S_{0,2}(t) + \varepsilon S_{1,2}(t). \end{aligned} \tag{2.18}$$

The general solution structure of (2.18) is (cf. for instance [12])

$$\begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} X_{0,2}(t) + \varepsilon X_{1,2}(t) + \dots \\ Y_{0,2}(t) + \varepsilon Y_{1,2}(t) + \dots \end{pmatrix} + \begin{pmatrix} \varepsilon m_{0,2}(t/\varepsilon) + \dots \\ n_{0,2}(t/\varepsilon) + \varepsilon n_{1,2}(t/\varepsilon) + \dots \end{pmatrix} \tag{2.19}$$

with smooth "outer solution terms" $X_{\ell,2}(t)$, $Y_{\ell,2}(t)$ and "inner solution terms" $m_{\ell,2}(t/\varepsilon)$, $n_{\ell,2}(t/\varepsilon)$ which are rapidly decaying away from $t = 0$ (smooth w.r.t. t/ε but not w.r.t. t). The derivatives w.r.t. t of the inner solution are affected by negative powers of ε. Recall that in our present context the goal is to choose *smooth* solutions of (2.18); to this end, special starting values have to be chosen such that a certain number of inner solution terms vanish. The purpose of this procedure is to ensure that those derivatives of $e_2(t)$ which appear as inhomogeneous terms within the other variational equations (defining $e_4(t), \dots, e_{2q}(t)$) and within the remainder equation (2.6) are $O(\varepsilon^0)$.

The construction of smooth solutions

$$\bar{e}_2(t) = \begin{pmatrix} x_2(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} X_{0,2}(t) + \varepsilon X_{1,2}(t) + \dots \\ Y_{0,2}(t) + \varepsilon Y_{1,2}(t) + \dots \end{pmatrix} \tag{2.20}$$

(with inner solution terms at a sufficiently high ε^p-level) is well known: Inserting (2.20) into (2.18) and equating coefficients of powers of ε leads to:

Coefficients of ε^0:

$$\begin{aligned} X_{0,2}'(t) &= c_1(t)X_{0,2}(t) + d_{11}(t)X_{0,2}(t) + d_{12}(t)Y_{0,2}(t) + R_{0,2}(t), \\ 0 &= -c_2(t)Y_{0,2}(t) + S_{0,2}(t). \end{aligned} \tag{2.21}$$

The second equation fixes $Y_{0,2}(t)$; the starting value is

$$Y_{0,2}(0) = \frac{S_{0,2}(0)}{c_2(0)}. \tag{2.22}$$

The starting value $X_{0,2}(0)$ for the first component $X_{0,2}(t)$ can be chosen arbitrarily.

This procedure continues in an obvious way: The $X_{\ell,2}(t)$ are solutions of certain smooth differential equations; the starting values $X_{\ell,2}(0)$ can be chosen arbitrarily. The $Y_{\ell,2}(t)$ are fixed in a recursive way by algebraic relations; the starting values $Y_{\ell,2}(0)$ are uniquely determined by the $Y_{k,2}(0)$ and the choice for the $X_{k,2}(0)$, $k < \ell$.

In the same way smooth solutions $e_4(t)$, $e_6(t)$, ... of the other variational equations can be constructed. The inhomogeneity of the equation defining $\bar{e}_{2i}(t) = T^{-1}(t)e_{2i}(t)$ depends recursively on certain derivatives of the $e_{2j}(t)$, $j < i$; for smooth $e_{2j}(t)$, $j < i$ it is obvious that the respective inhomogeneity is of a similar type as (2.17); it can be written in the form

$$v_{2i}(t) + \Lambda(t)w_{2i}(t) = \begin{pmatrix} R_{0,2i}(t) + \varepsilon R_{1,2i}(t) + \dots \\ \frac{1}{\varepsilon}S_{0,2i}(t) + S_{1,2i}(t) + \varepsilon S_{2,2i}(t) + \dots \end{pmatrix} \tag{2.23}$$

with certain smooth functions $R_{\ell,2i}(t)$ and $S_{\ell,2i}(t)$ which are defined in terms of the ε-expansions of the former $e_{2j}(t)$. Therefore smooth solutions are of the same form as (2.20):

$$\bar{e}_{2i}(t) = \begin{pmatrix} x_{2i}(t) \\ y_{2i}(t) \end{pmatrix} = \begin{pmatrix} X_{0,2i}(t) + \varepsilon X_{1,2i}(t) + \dots \\ Y_{0,2i}(t) + \varepsilon Y_{1,2i}(t) + \dots \end{pmatrix}; \tag{2.24}$$

the starting values for the first component can again be chosen arbitrarily (and the second components are again fixed by algebraic relations). Within our analysis for the remainder term R_ν (cf. sections 3 and 4) a *special choice* will be made for these starting values.

The techniqual details of the construction of smooth solutions are very similar as for the variational equations of the implicit Euler scheme (cf. [1], Lemma 3.1 and Proposition 3.2). In particular, the asymptotic correctness of the ε-expansions (2.20), (2.24) has to be verified, i.e., whether a remainder term in (2.20), (2.24) is indeed at a certain ε^p-level (cf. [1] for details).

Remark. The analogous construction has been used in [2]. But there is one significant difference: In [2] we have considered those special cases where the second components of the inhomogeneities (2.17), (2.23) turn out to be $O(\varepsilon^0)$ (and not $O(\varepsilon^{-1})$) despite the occurrence of $\Lambda(t)$. The obvious consequence is that, in this special case, the leading term $Y_{0,2i}(t)$ vanishes identically for all $e_{2i}(t)$. Therefore the choice $X_{0,2i}(0) = 0$ yields $e_{2i}(0) = O(\varepsilon)$. So, due to (2.12), we end up with $R_0 = O(h^2\varepsilon)$ which ensures the asymptotic correctness of (2.4) in the strongly stiff case $\varepsilon \le Ch^{2q}$.

The main consequence of the above construction is that for sufficiently smooth $e_{2i}(t)$ we have

$$I_\nu^\pm = O(h^{2q+2}), \qquad a_\nu = O(h^{2q+2}), \quad b_\nu = O(h^{2q+2}), \quad c_\nu = O(h^{2q+2}); \tag{2.25}$$

furthermore,

$$\|I_{\nu+1}^- - I_\nu^+\| = O(h^{2q+3}) \tag{2.26}$$

(recall the definition of $I_\nu^\pm$, a_ν, b_ν and c_ν at the begin of this section). Another consequence of the smoothness of the $e_{2i}(t)$ is that $u(t)$ from (2.10) is a moderate, smooth function. So the factor $h^2u(\hat{t}_\nu) + a_\nu$ within (2.11) is indeed $O(h^2)$ and therefore

$$h^2 G_\nu^{(1)} = O(h^2) \tag{2.27}$$

with a moderate O-constant. Moreover, by definition of G_ν,

$$G_\nu(0) = 0. \tag{2.28}$$

To show that the G_ν are Lipschitz continuous with a moderate Lipschitz constant, we may argue as follows: Recall that we have a priori assumed that the ζ_ν exist and are contained in a bounded neighbourhood of $z(t)$. From the fact that $\zeta_\nu - z(t_\nu)$ as well as the smooth functions $e_{2i}(t)$ are bounded, and from (2.4) we a priori know that the R_ν and consequently the actual arguments $R = \frac{1}{2}(R_{\nu-1} + R_\nu)$ of G_ν are bounded: $\|R\| \le C$. Furthermore, the arguments of f_{yy} in (2.11) are contained in a suitable bounded neighborhood of $z(t)$. Thus, due to our smoothness assumptions (2.3), we have

$$\|G_\nu(R)\| \le L_G\,\|R\| \qquad \text{for} \quad \|R\| \le C \tag{2.29}$$

with some moderate Lipschitz constant L_G (uniform w.r.t. ν).

The validity of (2.25)–(2.29) is essential for our structural analysis of the remainder term R_ν and will be used in sections 3 and 4 below.

The transformed remainder equation.

In sections 3 and 4 we shall present a structural analysis of the remainder term R_ν by means of discrete singular perturbation techniques. To this end we transform the remainder equation (2.6) analogously as we did for the variational equations (cf. (2.13), (2.14)). Some denotation:

$$\begin{aligned}
T_\nu &:= T(t_\nu), \quad T_{\nu-\frac12} := T(\hat t_\nu) = T(t_{\nu-\frac12}), \quad \Lambda_{\nu-\frac12} := \Lambda(\hat t_\nu) \\
\rho_\nu &:= T_\nu^{-1} R_\nu \\
\bar a_\nu &:= T_{\nu-\frac12}^{-1} a_\nu, \quad \bar b_\nu := T_{\nu-\frac12}^{-1} b_\nu, \quad \bar c_\nu := T_{\nu-\frac12}^{-1} c_\nu \\
\Theta_\nu^- &:= -T_{\nu-\frac12}^{-1}\,\frac{2}{h}\Big(T_\nu - T_{\nu-\frac12}\Big), \quad \Theta_{\nu-1}^+ := -T_{\nu-\frac12}^{-1}\,\frac{2}{h}\Big(T_{\nu-\frac12} - T_{\nu-1}\Big) \\
\Gamma_\nu(\rho) &:= T_{\nu-\frac12}^{-1} G_\nu(T_{\nu-\frac12}\rho)
\end{aligned} \tag{2.30}$$

According to (2.11), the smooth nonlinear functions $\Gamma_\nu(\rho)$ split in an obvious way into

$$\Gamma_\nu(\rho) = h^2\Gamma_\nu^{(1)}\cdot\rho + \Gamma_\nu^{(2)}(\rho)\cdot\rho^2. \tag{2.31}$$

Using the obvious relations

$$T_{\nu-\frac12}^{-1} R_{\nu-1} = \Big(I + \frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}, \quad T_{\nu-\frac12}^{-1} R_\nu = \Big(I - \frac{h}{2}\Theta_\nu^-\Big)\rho_\nu \tag{2.32}$$

and observing (2.1) we obtain from (2.6) after multiplication by $T_{\nu-\frac12}^{-1}$ the transformed remainder equation

$$\begin{aligned}
\frac1h\Big[\Big(I - \frac{h}{2}\Theta_\nu^-\Big)\rho_\nu - \Big(I + \frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}\Big] &= \Lambda_{\nu-\frac12}\,\frac12\Big[\Big(I + \frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1} + \Big(I - \frac{h}{2}\Theta_\nu^-\Big)\rho_\nu\Big] + \\
&\quad + \Gamma_\nu\bigg(\frac12\Big[\Big(I + \frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1} + \Big(I - \frac{h}{2}\Theta_\nu^-\Big)\rho_\nu\Big]\bigg) + \\
&\quad + \Lambda_{\nu-\frac12}\bar a_\nu + \bar b_\nu - \bar c_\nu.
\end{aligned} \tag{2.33}$$

The h-dependent quantities Θ_ν^- and $\Theta_{\nu-1}^+$, which are discrete analoga of $D(\hat t_\nu) = -T^{-1}(\hat t_\nu)T'(\hat t_\nu)$, are now expanded in powers of h. Using the denotation

$$D_\ell(t) := -\frac{1}{\ell}\,T^{-1}(t)\,\frac{d^\ell}{dt^\ell}T(t) \tag{2.34}$$

(with $D_1(t) = D(t)$, cf. (2.15)) we obtain by Taylor expansion about $\hat t_\nu$:

$$\begin{aligned}
\Theta_{\nu-1}^+ &= D_1(\hat t_\nu) - \frac{h}{2}D_2(\hat t_\nu) + \frac{h^2}{8}D_3(\hat t_\nu) - \frac{h^3}{48}D_4(\hat t_\nu) + \dots, \\
\Theta_\nu^- &= D_1(\hat t_\nu) + \frac{h}{2}D_2(\hat t_\nu) + \frac{h^2}{8}D_3(\hat t_\nu) + \frac{h^3}{48}D_4(\hat t_\nu) + \dots
\end{aligned} \tag{2.35}$$

Furthermore, the nonlinear functions $\Gamma_\nu(\rho) = T_{\nu-\frac12}^{-1}G_\nu(T_{\nu-\frac12}\rho)$, which depend on h^2 (cf. (2.11)), can be expanded into powers of h^2; using the splitting (2.31) we thus may write

$$\begin{aligned}
\Gamma_\nu(\rho) &= h^2\Gamma_\nu^{(1)}\cdot\rho + \Gamma_\nu^{(2)}(\rho)\cdot\rho^2 = \\
&= h^2\Big(\Gamma_{\nu,0}^{(1)} + h^2\Gamma_{\nu,2}^{(1)} + \dots\Big)\cdot\rho + \Big(\Gamma_{\nu,0}^{(2)}(\rho) + h^2\Gamma_{\nu,2}^{(2)}(\rho) + \dots\Big)\cdot\rho^2
\end{aligned} \tag{2.36}$$

and the $\Gamma^{(2)}_{\nu,2\ell}(\rho)$ are Taylor-expanded into

$$\Gamma^{(2)}_{\nu,2\ell}(\rho) = \Gamma^{(2)}_{\nu,2\ell}(0) + \frac{\partial}{\partial\rho}\Gamma^{(2)}_{\nu,2\ell}(0)\cdot\rho + \ldots \tag{2.37}$$

All these expansions are now inserted into the transformed remainder equation (2.33). After some simple manipulations, this leads us to the "expanded" remainder equation

$$\begin{aligned}
&\Big[I - \frac{h^2}{4}D_2(\hat{t}_\nu) - \frac{h^4}{96}D_4(\hat{t}_\nu) - \ldots\Big]\frac{1}{h}\Big(\rho_\nu - \rho_{\nu-1}\Big) - \\
&- \Big[D_1(\hat{t}_\nu) + \frac{h^2}{8}D_3(\hat{t}_\nu) + \ldots\Big]\frac{1}{2}\Big(\rho_{\nu-1} + \rho_\nu\Big) \;= \\
= \;& \Lambda_{\nu-\frac{1}{2}}\Big[I - \frac{h^2}{4}D_2(\hat{t}_\nu) - \frac{h^4}{96}D_4(\hat{t}_\nu) - \ldots\Big]\frac{1}{2}\Big(\rho_{\nu-1} + \rho_\nu\Big) - \\
&- \Lambda_{\nu-\frac{1}{2}}\frac{h^2}{4}\Big[D_1(\hat{t}_\nu) + \frac{h^2}{8}D_3(\hat{t}_\nu) + \ldots\Big]\frac{1}{h}\Big(\rho_\nu - \rho_{\nu-1}\Big) + \\
&+ h^2\Gamma^{(1)}_{\nu,0}\frac{1}{2}\Big(\rho_{\nu-1} + \rho_\nu\Big) - \frac{h^4}{4}\Gamma^{(1)}_{\nu,0}D_1(\hat{t}_\nu)\frac{1}{h}\Big(\rho_\nu - \rho_{\nu-1}\Big) + \ldots + \\
&+ \Gamma^{(2)}_{\nu,0}(0)\left[\frac{1}{2}\Big(\rho_{\nu-1} + \rho_\nu\Big) - \frac{h^2}{4}D_1(\hat{t}_\nu)\frac{1}{h}\Big(\rho_\nu - \rho_{\nu-1}\Big)\right]^2 + \ldots + \\
&+ \Lambda_{\nu-\frac{1}{2}}\bar{a}_\nu + \bar{b}_\nu - \bar{c}_\nu .
\end{aligned} \tag{2.38}$$

3 The Case $\varepsilon \leq C\,h^2$.

Let us now consider the case $\varepsilon \leq Ch^2$ where it will turn out that R_ν shows a very regular oscillating behavior (cf. the introductory discussion of section 1). In Theorem 3.1 below we shall give a quantitative description of R_ν by means of an expansion in powers of h^2. To keep the presentation transparent, Theorem 3.1 is restricted to an expansion up to the $O(h^6)$-level. Expansions up to a higher $O(h^{2q+2})$-level ($2q+2 = 8,\, 10,\, \ldots$) can be obtained in a very similar way; some technical modifications which are necessary for the general case will be discussed at the end of the present section.

Our analysis is based on a discrete "two-timing" singular perturbation technique which is a discrete modification of the two-timing (or multiscale) approach for the approximate solution of nonlinear differential equations of oscillatory type (cf. [11],[13]). We shall make an ansatz in powers h^2 involving terms of the form

$$\textit{smooth function}\,(t_\nu)\cdot(-1)^\nu \tag{3.1}$$

and show how to construct *special smooth solutions of the variational equations* (by means of choosing special starting values $e_{2i}(0)$) such that R_ν has indeed the desired structure. This leads us to a natural decomposition of the global error: The smooth terms are described by smooth solutions $e_{2i}(t)$ of the variational equations (2.5) whereas the non-smooth (oscillating) terms are represented by the h^2-expansion for R_ν.

To prove the asymptotic correctness of the h^2-expansion of R_ν, i.e. that the remainder term of this expansion is indeed at the desired $O(h^{2q+2})$-level (where $2q+2 = 6$ in Theorem 3.1) we shall need the following generalized stability estimate for nonlinear difference equations of the 'midpoint type' involving oscillating inhomogeneous terms at the reduced $O(h^{2q+1})$-level.

In the sequel, $\|\cdot\|$ denotes the Euclidean vector norm $\|\cdot\|_2$ as well as its associated matrix norm.

Lemma 3.1. *Consider a transformed difference equation of the type* (2.33)

$$\begin{aligned}
\frac{1}{h}\Big[\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu-\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}\Big] &= \Lambda_{\nu-\frac12}\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu\Big]+\\
&+\Phi_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu\Big]\Big)+\\
&+\Lambda_{\nu-\frac12}\bar a_\nu+\bar b_\nu-\bar c_\nu+ \qquad (3.2)\\
&+h^{2q+1}\delta_{\nu,1}(-1)^\nu+h^{2q+2}\delta_{\nu,2}+h^{2q+3}\Lambda_{\nu-\frac12}\,\delta_{\nu,3},\\
\eta_0 &= O(h^{2q+2})
\end{aligned}$$

with $\Lambda_{\nu-\frac12}$, $\Theta_\nu^\pm$ *and* $\bar a_\nu$, $\bar b_\nu$, $\bar c_\nu$ *as in* (2.33). *Assume that the nonlinear functions* $\Phi_\nu(\cdot)$ *satisfy* $\Phi_\nu(0)=0$ *and are uniformly Lipschitz continuous with a moderate Lipschitz constant. Let the* $\delta_{\nu,i}$ *be moderately sized and assume*

$$\|\delta_{\nu+1,1}-\delta_{\nu,1}\|=O(h) \qquad (3.3)$$

with a moderate O*-constant. Then, for* $\varepsilon\le Ch^2$,

$$\|\eta_\nu\|\le C_1(t_\nu)\|\eta_0\|+C_2(t_\nu)h^{2q+2} \qquad (3.4)$$

with certain moderate bounds $C_i(t)$. *Thus*

$$\|\eta_\nu\|=O(h^{2q+2}). \qquad (3.5)$$

Proof. According to the definition of $\bar a_\nu$ and $\bar c_\nu$ (cf. (2.30), (2.8)) we may write (analogously as in (2.32))

$$\begin{aligned}
\bar a_\nu &= \frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)T_{\nu-1}^{-1}I_\nu^-+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)T_\nu^{-1}I_\nu^+\Big],\\
\bar c_\nu &= \frac{1}{h}\Big[\Big(I-\frac{h}{2}\Theta_\nu^-\Big)T_\nu^{-1}I_\nu^+-\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)T_{\nu-1}^{-1}I_\nu^-\Big]
\end{aligned} \qquad (3.6)$$

Similarly as in Kraaijevanger [10] we introduce the auxiliary quantities

$$\eta_\nu^+ := \Big(I-\frac{h}{2}\Theta_\nu^-\Big)\Big(\eta_\nu+T_\nu^{-1}I_\nu^+\Big),\qquad \eta_{\nu-1}^- := \Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\Big(\eta_{\nu-1}+T_{\nu-1}^{-1}I_\nu^-\Big). \qquad (3.7)$$

With (3.7) and (2.30), the nonlinear term in (3.2) can be rewritten as

$$\begin{aligned}
\Phi_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu\Big]\Big) &= \Phi_\nu\Big(-T_{\nu-\frac12}^{-1}\frac{1}{2}\Big(I_\nu^-+I_\nu^+\Big)\Big)+ \qquad (3.8)\\
+\underbrace{\Big[\Phi_\nu\Big(\frac{1}{2}\Big(\eta_{\nu-1}^-+\eta_\nu^+\Big)-T_{\nu-\frac12}^{-1}\frac{1}{2}\Big(I_\nu^-+I_\nu^+\Big)\Big)-\Phi_\nu\Big(-T_{\nu-\frac12}^{-1}\frac{1}{2}\Big(I_\nu^-+I_\nu^+\Big)\Big)\Big]}_{=:\Psi_\nu\left(\frac12(\eta_{\nu-1}^-+\eta_\nu^+)\right)}
\end{aligned}$$

where, by assumption,

$$\Psi_\nu(0)=0,\qquad \|\Psi_\nu(\eta)\|\le L_\Phi\,\|\eta\|, \qquad (3.9)$$

L_Φ denoting a moderate Lipschitz constant for the $\Phi_\nu(\cdot)$.

With the denotation

$$j_\nu := \Phi_\nu\Big(-T^{-1}_{\nu-\frac12}\frac12\Big(I^-_\nu + I^+_\nu\Big)\Big) + \bar b_\nu + h^{2q+1}\delta_{\nu,1}(-1)^\nu + h^{2q+2}\delta_{\nu,2} + h^{2q+3}\Lambda_{\nu-\frac12}\delta_{\nu,3} \tag{3.10}$$

the difference equation (3.2) is equivalent to

$$\Big(I - \frac{h}{2}\Lambda_{\nu-\frac12}\Big)\eta^+_\nu = \Big(I + \frac{h}{2}\Lambda_{\nu-\frac12}\Big)\eta^-_{\nu-1} + h\,\Psi_\nu\Big(\frac12\Big(\eta^-_{\nu-1} + \eta^+_\nu\Big)\Big) + h\,j_\nu. \tag{3.11}$$

In the sequel we use the denotations

$$G_{\nu-\frac12} := \Big(I - \frac{h}{2}\Lambda_{\nu-\frac12}\Big)^{-1}\Big(I + \frac{h}{2}\Lambda_{\nu-\frac12}\Big), \qquad H_{\nu-\frac12} := \Big(I - \frac{h}{2}\Lambda_{\nu-\frac12}\Big)^{-1}. \tag{3.12}$$

Multiplication of (3.11) by $H_{\nu-\frac12}$ yields

$$\eta^+_\nu = G_{\nu-\frac12}\,\eta^-_{\nu-1} + h\,H_{\nu-\frac12}\,\Psi_\nu\Big(\frac12\Big(\eta^-_{\nu-1} + \eta^+_\nu\Big)\Big) + h\,H_{\nu-\frac12} j_\nu. \tag{3.13}$$

Denoting $\omega_0 := \frac{h}{2}\cdot\max\Big(\max_t \textit{spectral abscissa}\,(\Lambda(t)), 0\Big)$ (cf. (2.1)), $\omega := \omega_0 + \frac{h}{2}L_\Phi$, and using the inequalities

$$\begin{aligned} \|G_{\nu-\frac12}\| \le \frac{1+\omega_0}{1-\omega_0}, &\qquad \|H_{\nu-\frac12}\| \le \frac{1}{1-\omega_0}, \\ \|H_{\nu-\frac12}\,\tfrac{h}{2}\Lambda_{\nu-\frac12}\| = \|G_{\nu-\frac12} - H_{\nu-\frac12}\| &\le \frac{2+\omega_0}{1-\omega_0} \end{aligned} \tag{3.14}$$

we obtain after some simple manipulations

$$\|\eta^+_\nu\| \le \Big(\frac{1+\omega}{1-\omega}\Big)\,\|\eta^-_{\nu-1}\| + \frac{C}{1-\omega}h^{2q+2} \tag{3.15}$$

with some moderate constant C (cf. in particular (2.25)); (3.15) is valid under the mild stepsize restriction $\omega < 1$.

Note the occurrence of h^{2q+2} (but not h^{2q+3}) within (3.15) (caused by the term $h^{2q+1}\delta_{\nu,1}(-1)^\nu$ within j_ν). This is of course not sufficient to prove the desired result (3.4); it is necessary to consider two consecutive steps ($\nu-1 \to \nu \to \nu+1$) of (3.13). This leads us to

$$\begin{aligned} \eta^+_{\nu+1} \;=\; & G_{\nu+\frac12}G_{\nu-\frac12}\,\eta^-_{\nu-1} + h\,G_{\nu+\frac12}H_{\nu-\frac12}\Psi_\nu\Big(\frac12\Big(\eta^-_{\nu-1} + \eta^+_\nu\Big)\Big) + h\,G_{\nu+\frac12}H_{\nu-\frac12}j_\nu + \\ & + G_{\nu+\frac12}\Big(\eta^-_\nu - \eta^+_\nu\Big) + h\,H_{\nu+\frac12}\Psi_{\nu+1}\Big(\frac12\Big(\eta^-_\nu + \eta^+_{\nu+1}\Big)\Big) + h\,H_{\nu+\frac12}j_{\nu+1}; \end{aligned} \tag{3.16}$$

furthermore we express η^-_ν and $(\eta^-_\nu - \eta^+_\nu)$ by η^+_ν on the basis of (3.7):

$$\begin{aligned} \eta^-_\nu \;&=\; \Big(I + \frac{h}{2}\Theta^+_\nu\Big)\Big(I - \frac{h}{2}\Theta^-_\nu\Big)^{-1}\eta^+_\nu + \Big(I + \frac{h}{2}\Theta^+_\nu\Big)T^{-1}_\nu\Big(I^-_{\nu+1} - I^+_\nu\Big), \\ \eta^-_\nu - \eta^+_\nu \;&=\; \frac{h}{2}\Big(\Theta^-_\nu + \Theta^+_\nu\Big)\Big(I - \frac{h}{2}\Theta^-_\nu\Big)^{-1}\eta^+_\nu + \Big(I + \frac{h}{2}\Theta^+_\nu\Big)T^{-1}_\nu\Big(I^-_{\nu+1} - I^+_\nu\Big). \end{aligned} \tag{3.17}$$

Let ϑ denote a uniform bound for the $\|\Theta_\nu^\pm\|$, which is moderate due to our smoothness assumptions. Inserting (3.17) into (3.16) we obtain due to (3.9), (3.14) and due to $\|I+\frac{h}{2}\Theta_\nu^+\| \leq 1+\frac{h}{2}\vartheta$, $\|(I-\frac{h}{2}\Theta_\nu^-)^{-1}\| \leq 1/(1-\frac{h}{2}\vartheta)$:[6]

$$\begin{aligned}\|\eta_{\nu+1}^+\| \leq & \left(\frac{1+\omega_0}{1-\omega_0}\right)^2 \|\eta_{\nu-1}^-\| + \frac{1+\omega_0}{(1-\omega_0)^2}\,\frac{h}{2}L_\Phi\left[\|\eta_{\nu-1}^-\| + \|\eta_\nu^+\|\right] + \\ & + \frac{1+\omega_0}{1-\omega_0}\left[\frac{h\vartheta}{1-\frac{h}{2}\vartheta}\|\eta_\nu^+\| + \left(1+\frac{h}{2}\vartheta\right)\left\|T_\nu^{-1}(I_{\nu+1}^- - I_\nu^+)\right\|\right] + \\ & + \frac{1}{1-\omega_0}\,\frac{h}{2}L_\Phi\left[\frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\|\eta_\nu^+\| + \left(1+\frac{h}{2}\vartheta\right)\left\|T_\nu^{-1}(I_{\nu+1}^- - I_\nu^+)\right\| + \|\eta_{\nu+1}^+\|\right] + \\ & + h\left\|G_{\nu+\frac12}H_{\nu-\frac12}j_\nu + H_{\nu+\frac12}j_{\nu+1}\right\|.\end{aligned} \tag{3.18}$$

Now we proceed in the following way: On the left hand side we collect the terms involving $\|\eta_{\nu+1}^+\|$; on the right hand side we estimate $\|\eta_\nu^+\|$ by means of (3.15).[7] Then we multiply the resulting inequality by $(1-\omega_0)/(1-\omega)$ (which is positive under the stepsize restriction $\omega < 1$. The factor which then appears with $\|\eta_{\nu-1}^-\|$ on the right hand side can be rewritten as $((1+\omega)/(1-\omega))^2\,(1+\frac{h}{2}\vartheta)/(1-\frac{h}{2}\vartheta)$, and we obtain

$$\begin{aligned}\|\eta_{\nu+1}^+\| \leq & \left(\frac{1+\omega}{1-\omega}\right)^2 \frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\|\eta_{\nu-1}^-\| + \frac{1+\omega}{1-\omega}\left(1+\frac{h}{2}\vartheta\right)\left\|T_\nu^{-1}(I_{\nu+1}^- - I_\nu^+)\right\| + \\ & + \frac{1-\omega_0}{1-\omega}\,h\left\|G_{\nu+\frac12}H_{\nu-\frac12}j_\nu + H_{\nu+\frac12}j_{\nu+1}\right\| + O(h^{2q+3}).\end{aligned} \tag{3.19}$$

It has now to be shown that all inhomogeneous terms in this recursive estimate are $O(h^{2q+3})$. Indeed,

$$\left\|T_\nu^{-1}(I_{\nu+1}^- - I_\nu^+)\right\| = O(h^{2q+3}) \tag{3.20}$$

due to $\|T_\nu^{-1}\| = O(1)$ and due to (2.26). To show that $h\,\|G_{\nu+\frac12}H_{\nu-\frac12}j_\nu + H_{\nu+\frac12}j_{\nu+1}\|$ is also $O(h^{2q+3})$, we observe that the j_ν (cf. (3.10)) involve the following quantities:

(i) $\Phi_\nu(-T_{\nu-\frac12}^{-1}\frac12(I_\nu^- + I_\nu^+)) + \bar b_\nu + h^{2q+2}\delta_{\nu,2} = O(h^{2q+2})$ due to our smoothness assumptions and due to (2.25);

(ii) $h^{2q+3}\Lambda_{\nu-\frac12}\delta_{\nu,3}$ contains even a factor h^{2q+3} but is affected by $\Lambda_{\nu-\frac12}$;

(iii) $h^{2q+1}\delta_{\nu,1}(-1)^\nu$ is an oscillating term at the (reduced) $O(h^{2q+1})$-level.

Since $\|G_{\nu-\frac12}\|$, $\|H_{\nu-\frac12}\|$ and $\|H_{\nu-\frac12}\frac{h}{2}\Lambda_{\nu-\frac12}\|$ are moderate (cf. (3.14)), the contributions of the terms of type (i) and (ii) to $h\,\|G_{\nu+\frac12}H_{\nu-\frac12}j_\nu + H_{\nu+\frac12}j_{\nu+1}\|$ are $O(h^{2q+3})$, as required. The influence of the 'critical', oscillating terms of type (iii) has now to be studied carefully:

$$\begin{aligned}& h\,h^{2q+1}\left\|G_{\nu+\frac12}H_{\nu-\frac12}\delta_{\nu,1}(-1)^\nu + H_{\nu+\frac12}\delta_{\nu+1,1}(-1)^{\nu+1}\right\| = \\ & = h^{2q+2}\left\|G_{\nu+\frac12}H_{\nu-\frac12}\delta_{\nu,1} - H_{\nu+\frac12}\delta_{\nu+1,1}\right\| \leq \\ & \leq h^{2q+2}\left\|G_{\nu+\frac12}H_{\nu-\frac12} - H_{\nu+\frac12}\right\|\left\|\delta_{\nu,1}\right\| + h^{2q+2}\left\|H_{\nu+\frac12}\right\|\left\|\delta_{\nu+1,1} - \delta_{\nu,1}\right\|.\end{aligned} \tag{3.21}$$

[6]Note the mild stepsize restriction $\frac{h}{2}\vartheta < 1$.

[7]Note that in (3.18) all terms involving $\|\eta_\nu^+\|$ contain the factor h and therefore the inhomogeneous term in (3.15) leads to the $O(h^{2q+3})$-term in (3.19).

The second term on the right hand side of (3.21) is $O(h^{2q+3})$ due to (3.14) and assumption (3.3). To show that this is also true for the first term, the estimate $\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} - H_{\nu+\frac{1}{2}}\| = O(h)$ is required. By definition of $G_{\nu+\frac{1}{2}}$ and $H_{\nu\pm\frac{1}{2}}$,

$$G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} - H_{\nu+\frac{1}{2}} = \Big(I - \frac{h}{2}\Lambda_{\nu+\frac{1}{2}}\Big)^{-1}\left[\Big(I + \frac{h}{2}\Lambda_{\nu+\frac{1}{2}}\Big)\Big(I - \frac{h}{2}\Lambda_{\nu-\frac{1}{2}}\Big)^{-1} - I\right], \tag{3.22}$$

and the desired estimate

$$\Big\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} - H_{\nu+\frac{1}{2}}\Big\| = O(h) + O\Big(\frac{\varepsilon}{h}\Big) \quad \text{for } \varepsilon \le Ch^2 \tag{3.23}$$

follows easily by virtue of (2.2).

Thus, summarizing (3.19)–(3.23) we obtain the recursion

$$\|\eta^+_{\nu+1}\| \le \left(\frac{1+\omega}{1-\omega}\right)^2 \frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\|\eta^-_{\nu-1}\| + O(h^{2q+3}) \tag{3.24}$$

with a certain moderate O-constant. Estimating $\|\eta^-_{\nu-1}\|$ by $\|\eta^+_{\nu-1}\|$ on the basis of (3.17) and using once more (3.20) we end up with

$$\|\eta^+_{\nu+1}\| \le \left(\frac{1+\omega}{1-\omega}\right)^2 \left(\frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\right)^2\|\eta^+_{\nu-1}\| + O(h^{2q+3}). \tag{3.25}$$

Since $(1+\omega)/(1-\omega)\,(1+\frac{h}{2}\vartheta)/(1-\frac{h}{2}\vartheta) = 1 + O(h)$ with a moderate O-constant, a standard induction argument yields

$$\|\eta^+_\nu\| \le C^+_1(t_\nu)\|\eta^+_0\| + C^+_2(t_\nu)h^{2q+2} \tag{3.26}$$

with certain moderate bounds $C^+_i(t)$. Finally, the result (3.4) easily follows from (3.26) by definition of η^+_ν (cf. (3.7)) and due to (2.25). □

Theorem 3.1. *Consider a stiff problem* (1.3) *of the type* (2.1)–(2.3) *and assume* $\varepsilon \le Ch^2$. *Denote*

$$\chi := \frac{\varepsilon}{h^2}. \tag{3.27}$$

Let $[0, t_E]$ *be a subinterval of the whole integration interval with constant stepsize* h *and assume inductively that - after transformation by* $T^{-1}(t)$ *- the accumulated global error* $T^{-1}(0)(\zeta_0 - z(0))$ *at* $t = 0$ *satisfies*

$$T^{-1}(0)\Big(\zeta_0 - z(0)\Big) = h^2\begin{pmatrix} \tilde{x}_2 \\ \tilde{y}_2(\chi) \end{pmatrix} + h^4\begin{pmatrix} \tilde{x}_4(\chi) \\ \tilde{y}_4(\chi) \end{pmatrix} + O(h^6) \tag{3.28}$$

with a certain h*-independent quantity* $\tilde{x}_2$ *and with moderate quantities* $\tilde{y}_2(\chi)$, $\tilde{x}_4(\chi)$ *and* $\tilde{y}_4(\chi)$ *which depend smoothly on the parameter* χ. *Then the global error of the implicit midpoint rule on* $[0, t_E]$ *can be written as follows:*

$$\zeta_\nu - z(t_\nu) = h^2 e_2(t_\nu) + h^4 e_4(t_\nu; \chi) + R_\nu \tag{3.29}$$

where the h-independent function $e_2(t)$ and the function $e_4(t;\chi)$, which depends smoothly on the parameter χ, are smooth solutions of the variational equations (2.5). *Furthermore,*

$$R_\nu = T(t_\nu)\begin{pmatrix} h^4\alpha_4(t_\nu;\chi)(-1)^\nu \\ h^2\beta_2(t_\nu;\chi)(-1)^\nu + h^4\beta_4(t_\nu;\chi)(-1)^\nu \end{pmatrix} + O(h^6). \tag{3.30}$$

Here $\alpha_4(t;\chi)$, $\beta_2(t;\chi)$ and $\beta_4(t;\chi)$ are smooth functions in t which depend smoothly on the parameter χ. $\beta_2(t;\chi)$ is the solution of

$$\begin{aligned} \beta_2'(t;\chi) &= \left(d_{22}^{(1)}(t) - \frac{4\chi}{c_2(t)}\right)\beta_2(t;\chi), \\ \beta_2(0;\chi) &= \tilde{y}_2(\chi) - Y_{0,2}(0) \end{aligned} \tag{3.31}$$

(with $d_{22}^{(1)}(t)$ from (3.32) *below and where $Y_{0,2}(0)$ is the starting value of the leading term $Y_{0,2}(t)$ of the ε-expansion of the second component of $\bar{e}_2(t) = T^{-1}(t)e_2(t)$ - cf.* (2.20), (2.22)*). The equations defining $\alpha_4(t;\chi)$ and $\beta_4(t;\chi)$ are given in* (3.45) *and* (3.47) *below.*

Proof. We consider the "expanded" transformed remainder equation (2.38) and multiply the first component by h and the second component by ε/h. With the denotation

$$D_\ell(t) =: \begin{pmatrix} d_{11}^{(\ell)}(t) & d_{12}^{(\ell)}(t) \\ d_{21}^{(\ell)}(t) & d_{22}^{(\ell)}(t) \end{pmatrix} \tag{3.32}$$

this yields the following equations for the components x_ν and y_ν of the transformed remainder term $\rho_\nu = \begin{pmatrix} x_\nu \\ y_\nu \end{pmatrix}$:

$$\begin{aligned}
&\Big[1 - \frac{h^2}{4}d_{11}^{(2)}(\hat{t}_\nu) - \frac{h^4}{96}d_{11}^{(4)}(\hat{t}_\nu) + \dots\Big]\Big(x_\nu - x_{\nu-1}\Big) - \\
&-\Big[\quad \frac{h^2}{4}d_{12}^{(2)}(\hat{t}_\nu) + \frac{h^4}{96}d_{12}^{(4)}(\hat{t}_\nu) + \dots\Big]\Big(y_\nu - y_{\nu-1}\Big) - \\
&-h\Big[d_{11}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{11}^{(3)}(\hat{t}_\nu) + \dots\Big]\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big) - \\
&-h\Big[d_{12}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{12}^{(3)}(\hat{t}_\nu) + \dots\Big]\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big) \quad = \\
= \quad & h\,c_1(\hat{t}_\nu)\Big[1 - \frac{h^2}{4}d_{11}^{(2)}(\hat{t}_\nu) - \frac{h^4}{96}d_{11}^{(4)}(\hat{t}_\nu) + \dots\Big]\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big) - \\
&-h\,c_1(\hat{t}_\nu)\Big[\quad \frac{h^2}{4}d_{12}^{(2)}(\hat{t}_\nu) + \frac{h^4}{96}d_{12}^{(4)}(\hat{t}_\nu) + \dots\Big]\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big) - \\
&-c_1(\hat{t}_\nu)\,\frac{h^2}{4}\Big[d_{11}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{11}^{(3)}(\hat{t}_\nu) + \dots\Big]\Big(x_\nu - x_{\nu-1}\Big) - \\
&-c_1(\hat{t}_\nu)\,\frac{h^2}{4}\Big[d_{12}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{12}^{(3)}(\hat{t}_\nu) + \dots\Big]\Big(y_\nu - y_{\nu-1}\Big) + \\
&+h^2\,h\,\gamma(\hat{t}_\nu)\,\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big) + h^2\,h\,\gamma(\hat{t}_\nu)\,\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big) + \dots - \\
&-\frac{h^4}{4}\gamma(\hat{t}_\nu)\Big(x_\nu - x_{\nu-1}\Big) - \frac{h^4}{4}\gamma(\hat{t}_\nu)\Big(y_\nu - y_{\nu-1}\Big) + \dots + \\
&+h\cdot(\textit{higher order nonlinear terms}) + hc_1(\hat{t}_\nu)\bar{a}_{\nu,1} + h\bar{b}_{\nu,1} - h\bar{c}_{\nu,1}
\end{aligned} \tag{3.33}$$

and

$$
\begin{aligned}
& -\frac{\varepsilon}{h^2}\Big[\quad \frac{h^2}{4}d_{21}^{(2)}(\hat t_\nu)+\frac{h^4}{96}d_{21}^{(4)}(\hat t_\nu)+\dots\Big]\Big(x_\nu-x_{\nu-1}\Big)+\\
& +\frac{\varepsilon}{h^2}\Big[1-\frac{h^2}{4}d_{22}^{(2)}(\hat t_\nu)-\frac{h^4}{96}d_{22}^{(4)}(\hat t_\nu)+\dots\Big]\Big(y_\nu-y_{\nu-1}\Big)-\\
& -\frac{\varepsilon}{h^2}\,h\Big[d_{21}^{(1)}(\hat t_\nu)+\frac{h^2}{8}d_{21}^{(3)}(\hat t_\nu)+\dots\Big]\frac{1}{2}\Big(x_{\nu-1}+x_\nu\Big)-\\
& -\frac{\varepsilon}{h^2}\,h\Big[d_{22}^{(1)}(\hat t_\nu)+\frac{h^2}{8}d_{22}^{(3)}(\hat t_\nu)+\dots\Big]\frac{1}{2}\Big(y_{\nu-1}+y_\nu\Big)\quad=\\
= \quad & \frac{1}{h}\,c_2(\hat t_\nu)\Big[\quad \frac{h^2}{4}d_{21}^{(2)}(\hat t_\nu)+\frac{h^4}{96}d_{21}^{(4)}(\hat t_\nu)+\dots\Big]\frac{1}{2}\Big(x_{\nu-1}+x_\nu\Big)-\\
& -\frac{1}{h}\,c_2(\hat t_\nu)\Big[1-\frac{h^2}{4}d_{22}^{(2)}(\hat t_\nu)+\frac{h^4}{96}d_{22}^{(4)}(\hat t_\nu)+\dots\Big]\frac{1}{2}\Big(y_{\nu-1}+y_\nu\Big)+\\
& +c_2(\hat t_\nu)\,\frac{1}{4}\Big[d_{21}^{(1)}(\hat t_\nu)+\frac{h^2}{8}d_{21}^{(3)}(\hat t_\nu)+\dots\Big]\Big(x_\nu-x_{\nu-1}\Big)+\\
& +c_2(\hat t_\nu)\,\frac{1}{4}\Big[d_{22}^{(1)}(\hat t_\nu)+\frac{h^2}{8}d_{22}^{(3)}(\hat t_\nu)+\dots\Big]\Big(y_\nu-y_{\nu-1}\Big)+\\
& +h^2\,\frac{\varepsilon}{h^2}\,h\,\gamma(\hat t_\nu)\,\frac{1}{2}\Big(x_{\nu-1}+x_\nu\Big)+h^2\,\frac{\varepsilon}{h^2}\,h\,\gamma(\hat t_\nu)\,\frac{1}{2}\Big(y_{\nu-1}+y_\nu\Big)+\dots-\\
& -\frac{h^4}{4}\,\frac{\varepsilon}{h^2}\,\gamma(\hat t_\nu)\Big(x_\nu-x_{\nu-1}\Big)-\frac{h^4}{4}\,\frac{\varepsilon}{h^2}\,\gamma(\hat t_\nu)\Big(y_\nu-y_{\nu-1}\Big)+\dots+\\
& +h\frac{\varepsilon}{h^2}\cdot(\textit{higher order nonlinear terms})-\frac{1}{h}c_2(\hat t_\nu)\bar a_{\nu,2}+h\frac{\varepsilon}{h^2}\bar b_{\nu,2}-h\frac{\varepsilon}{h^2}\bar c_{\nu,2}
\end{aligned}
\tag{3.34}
$$

(Here $\gamma(t)$ is used as a generic denotation for the smooth functions originating from the Taylor expansions (2.36), (2.37).)

In the following any occurrence of ε is rewritten as $\varepsilon = h^2\chi$ where $\chi \le C$ is moderate under the present assumption $\varepsilon \le Ch^2$.

The starting condition to be fulfilled is (cf. (3.29))

$$
\rho_0 = T^{-1}(0)\Big(\zeta_0 - z(0)\Big) - h^2\bar e_2(0) - h^4\bar e_4(0;\chi), \tag{3.35}
$$

thus, due to our inductive assumption (3.28) and due to (2.24):

$$
\begin{aligned}
\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} &= h^2\begin{pmatrix} \tilde x_2 - X_{0,2}(0) - \varepsilon X_{1,2}(0) - \dots \\ \tilde y_2(\chi) - Y_{0,2}(0) - \varepsilon Y_{1,2}(0) - \dots \end{pmatrix} + h^4\begin{pmatrix} \tilde x_4(\chi) - X_{0,4}(0;\chi) - \dots \\ \tilde y_4(\chi) - Y_{0,4}(0;\chi) - \dots \end{pmatrix} + O(h^6)\\
&= h^2\begin{pmatrix} \tilde x_2 - X_{0,2}(0) \\ \tilde y_2(\chi) - Y_{0,2}(0) \end{pmatrix} + h^4\begin{pmatrix} \tilde x_4(\chi) - X_{0,4}(0;\chi) - \chi X_{1,2}(0) \\ \tilde y_4(\chi) - Y_{0,4}(0;\chi) - \chi Y_{1,2}(0) \end{pmatrix} + O(h^6).
\end{aligned}
\tag{3.36}
$$

Note that the quantities $X_{\ell,2}(0)$, $Y_{\ell,2}(0)$, $X_{\ell,4}(0;\chi)$ and $Y_{\ell,4}(0;\chi)$ are still undetermined. In (3.41) and (3.46)) below we shall make a special choice for the $X_{\ell,2}(0)$ and the $X_{\ell,4}(0;\chi)$. By construction of smooth solutions (cf. section 2) the $Y_{\ell,2}(0)$ and $Y_{\ell,4}(0;\chi)$ are then fixed. It will turn out that this choice can be made in such a way that - as indicated in the formulation of the theorem and in (3.36) - the function $e_2(t)$ does not depend on the parameter χ and is therefore h-independent. For $e_4(t;\chi)$, however, a χ-independent choice is not possible.[8]

[8]The discussion below will show that, actually, $Y_{0,4}(0;\chi) = Y_{0,4}(0)$ is χ-independent; but this is of minor relevance. The χ-dependence of $X_{0,4}$ is unavoidable.

For x_ν and y_ν we make the ansatz

$$\begin{aligned} x_\nu &= \qquad\qquad\qquad\quad h^4\alpha_4(t_\nu;\chi)(-1)^\nu + x_\nu^{(rem)} \\ y_\nu &= h^2\beta_2(t_\nu;\chi)(-1)^\nu + h^4\beta_4(t_\nu;\chi)(-1)^\nu + y_\nu^{(rem)} \end{aligned} \tag{3.37}$$

with certain functions $\alpha_{2i}(t;\chi)$ and $\beta_{2i}(t;\chi)$ which are to be determined. Before inserting into (3.33) and (3.34) we apply Taylor expansion about $\hat{t}_\nu$ to rewrite differences and arithmetic means of the terms $\alpha(t_\nu;\chi)(-1)^\nu$, $\beta(t_\nu;\chi)(-1)^\nu$ as

$$\begin{aligned} \alpha(t_\nu;\chi)(-1)^\nu - \alpha(t_{\nu-1};\chi)(-1)^{\nu-1} &= \Big(\alpha(t_{\nu-1};\chi) + \alpha(t_\nu;\chi)\Big)(-1)^\nu = \\ &= \left(2\alpha(\hat{t}_\nu;\chi) + \frac{h^2}{4}\alpha''(\hat{t}_\nu;\chi) + \dots\right)(-1)^\nu, \end{aligned} \tag{3.38}$$

$$\begin{aligned} \frac{1}{2}\Big(\alpha(t_{\nu-1};\chi)(-1)^{\nu-1} + \alpha(t_\nu;\chi)(-1)^\nu\Big) &= \frac{1}{2}\Big(\alpha(t_\nu;\chi) - \alpha(t_{\nu-1};\chi)\Big)(-1)^\nu = \\ &= h\left(\frac{1}{2}\alpha'(\hat{t}_\nu;\chi) + \frac{h^2}{48}\alpha'''(\hat{t}_\nu;\chi) + \dots\right)(-1)^\nu \end{aligned} \tag{3.39}$$

(analogously for the β's). After this preparation we equate coefficients of h^2, h^4, ... in (3.33) and (3.34), omitting the common factor $(-1)^\nu$:

- *Coefficients of* h^2:
  Since there is no h^2-term in the first component of (3.37), we have only an equation for β_2:
  $$2\chi\,\beta_2(\hat{t}_\nu;\chi) = -\frac{1}{2}c_2(\hat{t}_\nu)\,\beta_2'(\hat{t}_\nu;\chi) + \frac{1}{2}c_2(\hat{t}_\nu)d_{22}^{(1)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi) \tag{3.40}$$
  To satisfy the starting condition (3.36) at the h^2-level, the following choice has to be made:
  - In the first component we set
    $$X_{0,2}(0) := \tilde{x}_2 \tag{3.41}$$
    (recall that for a smooth solution $e_2(t)$ the $X_{\ell,2}(0)$ can be chosen arbitrarily - cf. section 2).
  - In the second component the $Y_{\ell,2}(0)$ are fixed by the requirement that $e_2(t)$ be smooth (cf. section 2). (In particular, $Y_{0,2}(0)$ is defined in terms of data functions - cf. (2.22).) So, due to (3.36), $\beta_2(0;\chi)$ is fixed by $\beta_2(0;\chi) = \tilde{y}_2(\chi) - Y_{0,2}(0)$.

  Thus we see from (3.40) that $\beta_2(t;\chi)$ is the solution of the initial value problem
  $$\begin{aligned} \beta_2'(t;\chi) &= \left(d_{22}^{(1)}(t) - \frac{4\,\chi}{c_2(t)}\right)\beta_2(t;\chi), \\ \beta_2(0;\chi) &= \tilde{y}_2(\chi) - Y_{0,2}(0) \end{aligned} \tag{3.42}$$
  which is smooth under the present assumption $\varepsilon \le Ch^2$ $(\chi \le C)$, depending smoothly on the parameter χ.

- *Coefficients of* h^4:

$$2\,\alpha_4(\hat{t}_\nu;\chi) - \frac{1}{2}d_{12}^{(2)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi) - \frac{1}{2}d_{12}^{(1)}(\hat{t}_\nu)\,\beta_2'(\hat{t}_\nu;\chi) = -\frac{1}{2}c_1(\hat{t}_\nu)d_{12}^{(1)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi) \quad (3.43)$$

$$\begin{aligned} &2\chi\,\beta_4(\hat{t}_\nu;\chi) + \frac{1}{4}\chi\,\beta_2''(\hat{t}_\nu;\chi) - \frac{1}{2}\chi d_{22}^{(1)}(\hat{t}_\nu)\,\beta_2'(\hat{t}_\nu;\chi) - \frac{1}{2}\chi d_{22}^{(2)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi) = \\ &= -\frac{1}{2}c_2(\hat{t}_\nu)\,\beta_4'(\hat{t}_\nu;\chi) - \frac{1}{48}c_2(\hat{t}_\nu)\,\beta_2'''(\hat{t}_\nu;\chi) + \frac{1}{2}c_2(\hat{t}_\nu)d_{21}^{(1)}(\hat{t}_\nu)\,\alpha_4(\hat{t}_\nu;\chi) + \\ &+ \frac{1}{2}c_2(\hat{t}_\nu)d_{22}^{(1)}(\hat{t}_\nu)\,\beta_4(\hat{t}_\nu;\chi) + \frac{1}{16}c_2(\hat{t}_\nu)d_{22}^{(1)}(\hat{t}_\nu)\,\beta_2''(\hat{t}_\nu;\chi) + \\ &+ \frac{1}{8}c_2(\hat{t}_\nu)d_{22}^{(2)}(\hat{t}_\nu)\,\beta_2'(\hat{t}_\nu;\chi) + \frac{1}{16}c_2(\hat{t}_\nu)d_{22}^{(3)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi). \end{aligned} \quad (3.44)$$

(3.43) is an *algebraic* equation fixing $\alpha_4(t;\chi)$. Expressing $\beta_2'(t;\chi)$ by means of (3.42) we obtain

$$\alpha_4(t;\chi) = \frac{1}{4}\left[d_{12}^{(1)}(t)\left(d_{22}^{(1)}(t) - c_1(t) - \frac{4\,\chi}{c_2(t)}\right) + d_{12}^{(2)}(t)\right]\beta_2(t;\chi). \quad (3.45)$$

Thus also $\alpha_4(0;\chi)$ is fixed; to satisfy the starting condition (3.36) at the h^4-level in the first component we make use of the freedom of choosing $X_{0,4}(0;\chi)$ and $X_{1,2}(0)$ (recall again that smooth solutions e_2, e_4 exist for an arbitrary choice of the $X_{\ell,2}(0)$, $X_{\ell,4}(0)$). Our choice is

$$\begin{aligned} X_{1,2}(0) &:= 0, \\ X_{0,4}(0;\chi) &:= \alpha_4(0;\chi) - \tilde{x}_4(\chi), \end{aligned} \quad (3.46)$$

where $\alpha_4(0;\chi)$ is determined by (3.45).

Note that $X_{0,2}(0)$ (cf. (3.41)) and $X_{1,2}(0)$ (cf. (3.46)) - and therefore also $Y_{1,2}(0)$ (which depends on $X_{0,2}(0)$ and $X_{1,2}(0)$) - are independent of χ. As a consequence, $e_2(t)$ is independent of χ as asserted in the formulation of the theorem.[9] (Within the ε-expansion (2.20) of $e_2(t)$, the $X_{\ell,2}(0)$, $\ell \geq 2$ can be chosen arbitrarily because this choice does only influence the $O(h^6)$-remainder term.)

The starting value for the second component is fixed by (3.36): $\beta_4(0;\chi) = \tilde{y}_4(\chi) - Y_{0,4}(0) - \chi Y_{1,2}(0)$. (Recall that $Y_{0,4}(0)$ and $Y_{1,2}(0)$ are fixed by the requirement that $e_2(t)$ and $e_4(t)$ be smooth; they can be expressed in terms of data functions and of already known $X_{\ell,k}(0)$, $Y_{\ell,k}(0)$.) From (3.44) we see that $\beta_4(t;\chi)$ is the solution of the smooth initial value problem

$$\begin{aligned} \beta_4'(t;\chi) &= \left(d_{22}^{(1)}(t) - \frac{4\,\chi}{c_2(t)}\right)\beta_4(t;\chi) + j_4(t;\chi), \\ \beta_4(0;\chi) &= \tilde{y}_4(\chi) - Y_{0,4}(0) - \chi Y_{1,2}(0) \end{aligned} \quad (3.47)$$

where the inhomogeneity $j_4(t;\chi)$ depends recursively on $\beta_2(t;\chi)$ and $\alpha_4(t;\chi)$; using (3.42) and (3.45) we can write this as

$$j_4(t;\chi) = \Big(u_0(t) + \chi u_1(t) + \chi^2 u_2(t) + \chi^3 u_3(t)\Big)\beta_2(t;\chi) \quad (3.48)$$

where the $u_\ell(t)$ are defined in terms of data functions. So it is obvious that also $\beta_4(t;\chi)$ depends smoothly on the parameter χ.

---

[9] A consequence of this is that also $Y_{0,4}(0)$, which depends only on data functions and on $e_2(t)$ (cf. (2.5)), is independent of χ. But this is not true for the $Y_{\ell,4}(0)$, $\ell \geq 1$ which are influenced by $X_{0,4}(0;\chi)$.

- *Estimation of the h^6-remainder term:*
To estimate the remainder term (cf. (3.37))

$$\rho_\nu^{(rem)} = \begin{pmatrix} x_\nu^{(rem)} \\ y_\nu^{(rem)} \end{pmatrix} \tag{3.49}$$

we collect all terms in (3.33) and (3.34) (with (3.37) substituted for x_ν and y_ν) which have not yet appeared so far. Then we re-scale the resulting equations for $x_\nu^{(rem)}$ and $y_\nu^{(rem)}$ (i.e., these equations are multiplied by $1/h$ and h/ε, respectively). In order to apply Lemma 3.1 we observe that the resulting difference equation for $\rho_\nu^{(rem)}$ is of the type (3.2). The nonlinear term $\Phi_\nu\Big(\frac{1}{2}\Big[\big(I+\frac{h}{2}\Theta_{\nu-1}^+\big)\rho_{\nu-1}^{(rem)} + \big(I-\frac{h}{2}\Theta_\nu^-\big)\rho_\nu^{(rem)}\Big]\Big)$ in this difference equation is defined in the following way: According to the h^2-expansion of ρ_ν,

$$\rho_\nu = \underbrace{\begin{pmatrix} h^4\alpha_4(t_\nu;\chi)(-1)^\nu \\ h^2\beta_2(t_\nu;\chi)(-1)^\nu + h^4\beta_4(t_\nu;\chi)(-1)^\nu \end{pmatrix}}_{=:\rho_\nu^{(exp)}} + \rho_\nu^{(rem)}, \tag{3.50}$$

where $\rho_\nu^{(exp)}$ is already determined and where $\rho_\nu^{(rem)}$ remains to be determined, we re-write the original nonlinear term $\Gamma_\nu(\cdot)$ in (2.33) as

$$\begin{aligned} &\Gamma_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1} + \Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu\Big]\Big) \equiv \\ \equiv\ &\Gamma_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)(\rho_{\nu-1}^{(exp)}+\rho_{\nu-1}^{(rem)}) + \Big(I-\frac{h}{2}\Theta_\nu^-\Big)(\rho_\nu^{(exp)}+\rho_\nu^{(rem)})\Big]\Big) \equiv \\ \equiv\ &\Gamma_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}^{(exp)} + \Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu^{(exp)}\Big]\Big) + \\ &+\Phi_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}^{(rem)} + \Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu^{(rem)}\Big]\Big). \end{aligned} \tag{3.51}$$

From this implicit definition of Φ_ν and due to our smoothness assumptions w.r.t. Γ_ν it is obvious that Φ_ν satisfies $\Phi_\nu(0)=0$ and is Lipschitz continuous with some moderate Lipschitz constant, as assumed in Lemma 3.1. The first term on the right hand side of (3.51) depends only on already specified quantities and appears as an inhomogeneous term within the equation for $\rho_\nu^{(rem)}$ (cf. (iii) below).
To enable the application of Lemma 3.1 it remains to be shown that the inhomogeneity of the difference equation for $\rho_\nu^{(rem)}$ is indeed of the form as assumed in (3.2) (with $2q+2=6$). Besides $\Lambda_{\nu-\frac{1}{2}}\bar{a}_\nu + \bar{b}_\nu - \bar{c}_\nu$, the following terms appear:

(i) Terms originating from the left hand side of (3.33) and (3.34), resp., yield inhomogeneous terms of the type $h^5\delta_{\nu,1}(-1)^\nu$. To see this, consider for instance

$$-\frac{h^2}{4}d_{11}^{(2)}(\hat{t}_\nu)\Big(x_\nu - x_{\nu-1}\Big) \tag{3.52}$$

occurring on the left hand of (3.33). After re-scaling of (3.33) by $1/h$ this yields the inhomogeneous term

$$\begin{aligned} &-\frac{1}{h}\frac{h^2}{4}d_{11}^{(2)}(\hat{t}_\nu)\,h^4\Big(\alpha_4(t_\nu;\chi)(-1)^\nu - \alpha_4(t_{\nu-1};\chi)(-1)^{\nu-1}\Big) = \\ &= -\frac{h^5}{2}d_{11}^{(2)}(\hat{t}_\nu)\,\alpha_4'(\hat{t}_\nu;\chi)(-1)^\nu + O(h^7) \end{aligned} \tag{3.53}$$

(cf. (3.38)). Another term which occurs on the left hand side of (3.34) is

$$-\chi\, h\, d_{21}^{(1)}(\hat{t}_\nu)\frac{1}{2}\Big(x_{\nu-1}+x_\nu\Big); \tag{3.54}$$

after re-scaling of (3.34) by $h/\varepsilon = 1/(h\chi)$ this yields the inhomogeneous term

$$\begin{aligned}-\frac{1}{\chi h}\,\chi h\, d_{21}^{(1)}(\hat{t}_\nu)\, h^4\,\frac{1}{2}\Big(\alpha_4(t_{\nu-1};\chi)(-1)^{\nu-1}+\alpha_4(t_\nu;\chi)(-1)^\nu\Big) &= \\ = -\frac{h^5}{2} d_{21}^{(1)}(\hat{t}_\nu)\,\alpha_4(\hat{t}_\nu;\chi)(-1)^\nu + O(h^7)\end{aligned} \tag{3.55}$$

(cf. (3.39)). Collecting all these terms indeed yields an inhomogeneous term of the type $h^5\delta_{\nu,1}(-1)^\nu$ where $\delta_{\nu,1}$ satisfies (3.3) due to the smoothness of $\alpha_4(t;\chi)$, ...

(ii) For those terms on the right hand side of (3.33) and (3.34) which contain the factor $c_1(\hat{t}_\nu)$ or $c_2(\hat{t}_\nu)$, respectively, similar considerations show that - after re-scaling - there occur h^5, h^7, ... - terms in the first and $\frac{h^7}{\varepsilon}$, $\frac{h^9}{\varepsilon}$, ... - terms in the second component. The h^5-terms from the first component contribute to $h^5\delta_{\nu,1}(-1)^\nu$; the other terms can be recombined into

$$h^7\Lambda_{\nu-\frac{1}{2}}\delta_{\nu,3} \tag{3.56}$$

with some moderate quantity $\delta_{\nu,3}$.

(iii) A further inhomogeneous term is $\Gamma_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}^{(exp)}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu^{(exp)}\Big]\Big)$ (cf. (3.51); note that the nonlinear terms gave no contribution at the h^2- or h^4-levels). According to the splitting (2.31) there occur “linear” terms originating from $h^2\Gamma_\nu^{(1)}(\cdots)$ (cf. the terms involving the generic factor $\gamma(\hat{t}_\nu)$ in (3.33) and (3.34)) and “quadratic” terms originating from $\Gamma_\nu^{(2)}(\cdots)\cdot(\cdots)^2$. Analogous considerations as in (ii) above show that the linear terms contribute to $h^5\delta_{\nu,1}(-1)^\nu$. The quadratic terms are at least $O(h^6)$ due to $\Gamma_\nu^{(2)}(0)=0$ and $\|\Gamma_\nu^{(2)}(\rho)\|\le L_\Gamma\|\rho\|$ and because

$$\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}^{(exp)}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu^{(exp)}\Big]=O(h^6). \tag{3.57}$$

(Note that, due to (3.50), the leading term within $\Big\{\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\rho_{\nu-1}^{(exp)}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\rho_\nu^{(exp)}\Big]\Big\}^2$ involves $\frac{1}{2}\Big(h^2\beta_2(t_{\nu-1};\chi)(-1)^{\nu-1}+h^2\beta_2(t_\nu;\chi)(-1)^\nu\Big)^2$ which is indeed $O(h^6)$ due to (3.39).)

Since $\rho_0^{(rem)}=O(h^6)$ by construction, the desired estimate $\rho_\nu^{(rem)}=O(h^6)$ follows now by application of Lemma 3.1. □

Theorem 3.1 can be applied inductively to describe the error structure for *nonequidistant* grids: Consider the endpoint $t=t_E=\nu_E\, h$ of the current subinterval. Due to (2.24), (3.29) and (3.30) we see that for $\varepsilon\le Ch^2$ the transformed global error at $t=t_E$ can be written as (we assume ν_E even)

$$\begin{aligned}&\tilde{h}^2\Big(\frac{h}{\tilde{h}}\Big)^2\left[\begin{pmatrix}X_{0,2}(t_E)\\ Y_{0,2}(t_E)\end{pmatrix}+\begin{pmatrix}0\\ \beta_2(t_E;\chi)\end{pmatrix}\right]+\\ &+\ \tilde{h}^4\Big(\frac{h}{\tilde{h}}\Big)^4\left[\begin{pmatrix}X_{0,4}(t_E;\chi)+\chi X_{1,2}(t_E)\\ Y_{0,4}(t_E)\ +\ \chi Y_{1,2}(t_E)\end{pmatrix}+\begin{pmatrix}\alpha_4(t_E;\chi)\\ \beta_4(t_E;\chi)\end{pmatrix}\right]+O(\tilde{h}^6)\end{aligned} \tag{3.58}$$

where $\tilde{h}$ denotes the stepsize in the subinterval following $[0, t_E]$. So the global error at $t = t_E$ has the same structure as assumed in (3.28), and the induction proceeds to the next subinterval (note that the "new" $\tilde{x}_2$ is again χ-independent).

The begin of this induction is trivial if, in the very first subinterval, the integration starts on a smooth solution of the original problem. If there is a transient phase, however, the above argumentation does not work (because smooth solutions of the variational equations do not exist in the transients). However, (3.28) can also be assumed immediately after the transient phase: Since the B-convergence theory shows that the global error of the midpoint rule is (essentially) proportional to the length of the integration interval (if the integration starts on the true solution) and since the length of the transient phase is always $O(\varepsilon)$, the error bound immediately after the transients contains a factor ε. If one assumes further that the stepsizes in the transients are adjusted to the gradually improving smoothness of the solution such that local errors are approximately equidistributed, then the error after the transients is $O(\varepsilon h^2) = O(\chi h^4)$ (where h denotes the stepsize *after* the transient phase). So - immediately after the transients - (3.28) is valid with $\tilde{x}_2 = \tilde{y}_2(\chi) = 0$ and with certain moderate quantities $\tilde{x}_4(\chi)$ and $\tilde{y}_4(\chi)$.

We note that in the case where in the very first subinterval the integration starts on a smooth solution $z(0)$, the above proof can be modified in such a way that, in this first subinterval, both $e_2(t)$ *and* $e_4(t)$ are independent of χ, i.e., independent of h. This is possible because in this case the quantities $\tilde{x}_2, \ldots, \tilde{y}_4$ from (3.28) are zero; consequently $\beta_2(0)$ in (3.42) is independent of χ and $\alpha_4(0;\chi)$ in (3.45) is a polynomial of degree 1 in χ. So - instead of the above choice (3.46) - the choice

$$\begin{aligned} X_{1,2}(0) &:= -\frac{d_{12}^{(1)}(0)}{c_2(0)}\,\beta_2(0), \\ X_{0,4}(0) &:= \frac{1}{4}\left[d_{12}^{(1)}(0)\left(d_{22}^{(1)}(0) - c_1(0)\right) + d_{12}^{(2)}(0)\right]\beta_2(0) \end{aligned} \tag{3.59}$$

would indeed guarantee that $e_4(t)$ is also h-independent. But this is of minor practical relevance because, after a change of stepsize, the starting values for the new subinterval are influenced by the χ-dependent $O(h^2)$-remainder term (3.30) such that in the following subinterval e_4 is again χ-dependent.

Again we recall that the special problem class where $d_{21}^{(1)}(t) = O(\varepsilon)$ (weak coupling from the non-stiff to the stiff component), and where the right hand side $f(t,y)$ of (1.3) satisfies certain additional smoothness assumptions, has already been analyzed in [2]. For this problem class the structure of smooth solutions of the transformed variational equations is

$$\bar{e}_{2i}(t) = \begin{pmatrix} X_{0,2i}(t) + \varepsilon X_{1,2i}(t) + \ldots \\ \varepsilon Y_{1,2i}(t) + \ldots \end{pmatrix} \tag{3.60}$$

(i.e., $Y_{0,2i}(t) \equiv 0$); Theorem 3.1 of [2] shows that a full asymptotic expansion exists in this case under the assumption $\varepsilon \le Ch^{2q}$. In our present terminology one may say that, for $\varepsilon \le Ch^{2q}$, the quantities $\alpha_{2i}(t;\chi)$ and $\beta_{2i}(t;\chi)$ vanish identically up to the $O(h^{2q+2})$-level (which can be seen by a straightforward modification of the arguments used in the proof of Theorem 3.1 above).

We conclude this section by some remarks concerning *longer expansions* than covered by Theorem 3.1. The basic procedure is the same as outlined in the proof of Theorem 3.1. In the nonlinear case, however, the remainder term R_ν will also contain smooth, non-oscillating terms at the h^6-level. This can be explained as follows: Recall that within the equations for the $\alpha_{2i}(t;\chi)$ and $\beta_{2i}(t;\chi)$ there occur inhomogeneous terms which depend recursively on α's and β's which are already determined. At the h^2- and h^4-levels this dependence is purely linear and therefore all terms have the factor $(-1)^\nu$ in common. (We omitted this common factor in (3.40), (3.43) and (3.44) above.) However, beginning with the h^6-level there occur also inhomogeneous terms which depend in a *nonlinear* way on the former α's and β's. But an even power of an oscillating term with a smooth amplitude is a smooth function, as for instance

$$\Big[\beta_2(\hat{t}_\nu;\chi)(-1)^\nu\Big]^2 = \beta_2^2(\hat{t}_\nu;\chi). \tag{3.61}$$

The techniqual consequence is that, within the singular perturbation analysis of R_ν, "smooth variables" $A_{2i}(t;\chi)$, $B_{2i}(t;\chi)$ have to be introduced in addition to the "oscillating variables" $\alpha_{2i}(t;\chi)$ and $\beta_{2i}(t;\chi)$:

$$\begin{aligned} x_\nu &= \phantom{h^2\beta_2(t_\nu;\chi)(-1)^\nu + {}} h^4\alpha_4(t_\nu;\chi)(-1)^\nu + h^6\alpha_6(t_\nu;\chi)(-1)^\nu + h^6 A_6(t_\nu;\chi) + \dots \\ y_\nu &= h^2\beta_2(t_\nu;\chi)(-1)^\nu + h^4\beta_4(t_\nu;\chi)(-1)^\nu + h^6\beta_6(t_\nu;\chi)(-1)^\nu + h^6 B_6(t_\nu;\chi) + \dots \end{aligned} \tag{3.62}$$

For the oscillating variables we use again the Taylor expansions (3.38), (3.39); the smooth variables are expanded according to

$$\begin{aligned} A(t_\nu;\chi) - A(t_{\nu-1};\chi) &= h\,A'(\hat{t}_\nu;\chi) + \frac{h^3}{24}A'''(\hat{t}_\nu;\chi) + \dots, \\ \frac{1}{2}\Big(A(t_{\nu-1};\chi) + A(t_\nu;\chi)\Big) &= A(\hat{t}_\nu;\chi) + \frac{h^2}{8}A''(\hat{t}_\nu;\chi) + \dots \end{aligned} \tag{3.63}$$

(analogously for the B's). Inserting into (3.33), (3.34) and equating coefficients of powers of h yields equations for the smooth and for the oscillating variables. It turns out that the $B_{2i}(t;\chi)$ are determined by algebraic equations whereas the $A_{2i}(t;\chi)$ are solutions of certain smooth initial value problems. (In particular, $B_6(t;\chi)$ results to zero.)

Let us summarize the essential result of all these considerations: For $\varepsilon \le Ch^2$ the (transformed) global error of the implicit midpoint rule splits up into smooth terms (described by smooth solutions of the variational equations and by the $h^{2i}A_{2i}(t;\chi)$ and $h^{2i}B_{2i}(t;\chi)$, $i \ge 3$) and oscillating terms $h^{2i}\alpha_{2i}(t;\chi)(-1)^\nu$ and $h^{2i}\beta_{2i}(t;\chi)(-1)^\nu$ (with smooth amplitude functions). Moreover, since $\chi = \varepsilon/h^2$, the global error for a fixed $t_\nu = \nu h$, ν even, is an even function in h. The degree of smoothness w.r.t. h^2 (which is, for instance, relevant for a theoretical discussion of extrapolation) can be described quantitatively by differentiating the various error terms w.r.t. h^2. This smoothness increases as the ratio ε/h^2 decreases.

4 The Case $\varepsilon \le C\,h$.

Under the weaker assumption $\varepsilon \le Ch$ the nonsmooth terms of the global error occurring in R_ν cannot be represented by 'purely oscillating' terms (3.1). Unless even $\varepsilon \le Ch^2$ holds, the oscillation is superposed by a certain damping behavior characterized by terms of the type

$$\textit{smooth function}\,(t_\nu)\cdot\Pi_\nu \tag{4.1}$$

where

$$\Pi_\nu := \prod_{\ell=1}^{\nu}\left(\frac{1-\frac{c_2(\hat{t}_\ell)h}{2\varepsilon}}{1+\frac{c_2(\hat{t}_\ell)h}{2\varepsilon}}\right). \tag{4.2}$$

Note that for decreasing ε the Π_ν tend towards $(-1)^\nu$, i.e. the damping becomes weaker and the pure oscillation (discussed in section 3) reappears.

In Theorem 4.1 below we shall give a quantitative description of R_ν by means of an expansion in powers of h^2, involving terms of the type (4.1). Theorem 4.1 is restricted to an expansion up to the $O(h^4)$-level; expansions up to a higher $O(h^{2q+2})$-level (2q+2=6, 8, ...) are slightly more complicated but can be obtained by similar techniques.

For the proof of the asymptotic correctness of the expansion derived in Theorem 4.1 we shall need the following stability lemma.

Lemma 4.1. *Consider a transformed difference equation of the type* (2.33)

$$\begin{aligned} \frac{1}{h}\Big[\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu-\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}\Big] &= \Lambda_{\nu-\frac{1}{2}}\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu\Big]+ \\ &+\Phi_\nu\Big(\frac{1}{2}\Big[\Big(I+\frac{h}{2}\Theta_{\nu-1}^+\Big)\eta_{\nu-1}+\Big(I-\frac{h}{2}\Theta_\nu^-\Big)\eta_\nu\Big]\Big)+ \\ &+\Lambda_{\nu-\frac{1}{2}}\bar a_\nu+\bar b_\nu-\bar c_\nu+ \qquad (4.3) \\ &+h^{2q+2}\delta_\nu+\begin{pmatrix} h^{2q+1}\sigma_\nu\Pi_{\nu-1} \\ \frac{h^{2q+3}}{\varepsilon}\tau_\nu\Pi_{\nu-1}\end{pmatrix}, \\ \eta_0 &= O(h^{2q+2}) \end{aligned}$$

with $\Lambda_{\nu-\frac{1}{2}}$, $\Theta_\nu^\pm$ *and* $\bar a_\nu$, $\bar b_\nu$, $\bar c_\nu$ *as in* (2.33). *Assume that the nonlinear functions* $\Phi_\nu(\cdot)$ *satisfy* $\Phi_\nu(0)=0$ *and are uniformly Lipschitz continuous with a moderate Lipschitz constant. Let* δ_ν, σ_ν and τ_ν *be moderately sized and assume*

$$|\sigma_{\nu+1}-\sigma_\nu|=O(h) \qquad (4.4)$$

with a moderate O*-constant. Then, for* $\varepsilon\le Ch$,

$$\|\eta_\nu\|\le C_1(t_\nu)\|\eta_0\|+C_2(t_\nu)h^{2q+2} \qquad (4.5)$$

with certain moderate bounds $C_i(t)$. *Thus*

$$\|\eta_\nu\|=O(h^{2q+2}). \qquad (4.6)$$

Proof. The proof if similar to that of Lemma 3.1. (3.6)–(3.20) remain valid except that, instead of (3.10), we now have

$$j_\nu := \Phi_\nu\Big(-T_{\nu-\frac{1}{2}}^{-1}\frac{1}{2}\Big(I_\nu^-+I_\nu^+\Big)\Big)+\bar b_\nu+h^{2q+2}\delta_\nu+\begin{pmatrix} h^{2q+1}\sigma_\nu\Pi_{\nu-1} \\ \frac{h^{2q+3}}{\varepsilon}\tau_\nu\Pi_{\nu-1}\end{pmatrix}. \qquad (4.7)$$

Consider the recursion (3.19) with j_ν from (4.7). To show that $h\,\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}}j_\nu+H_{\nu+\frac{1}{2}}j_{\nu+1}\|=O(h^{2q+3})$, as required, we observe that the j_ν involve

(i) terms which are $O(h^{2q+2})$,

(ii) 'critical' terms of the form

$$\begin{pmatrix} h^{2q+1}\sigma_\nu\Pi_{\nu-1} \\ \frac{h^{2q+3}}{\varepsilon}\tau_\nu\Pi_{\nu-1}\end{pmatrix} =: \bar j_\nu\,\Pi_{\nu-1}. \qquad (4.8)$$

Since $\|G_{\nu-\frac{1}{2}}\|$ and $\|H_{\nu-\frac{1}{2}}\|$ are moderate (cf. (3.14)), the contribution of the terms of type (i) to $h\,\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}}j_\nu + H_{\nu+\frac{1}{2}}j_{\nu+1}\|$ is $O(h^{2q+3})$. The influence of the 'critical' terms (4.8) has now to be studied carefully. With

$$Q_\nu := \frac{1-\frac{c_2(\hat{t}_\nu)h}{2\varepsilon}}{1+\frac{c_2(\hat{t}_\nu)h}{2\varepsilon}} \tag{4.9}$$

we have

$$\begin{aligned} & h\Big\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}}\bar{\jmath}_\nu\,\Pi_{\nu-1} + H_{\nu+\frac{1}{2}}\bar{\jmath}_{\nu+1}\Pi_\nu\Big\| \le \\ \le\; & h\Big\|G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}}\bar{\jmath}_\nu + Q_\nu H_{\nu+\frac{1}{2}}\bar{\jmath}_{\nu+1}\Big\||\Pi_{\nu-1}| \le \\ \le\; & h\Big\|\Big(G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} + Q_\nu H_{\nu+\frac{1}{2}}\Big)\bar{\jmath}_\nu\Big\||\Pi_{\nu-1}| + h|Q_\nu|\Big\|H_{\nu+\frac{1}{2}}\Big(\bar{\jmath}_{\nu+1}-\bar{\jmath}_\nu\Big)\Big\||\Pi_{\nu-1}|. \end{aligned} \tag{4.10}$$

Due to (2.2) we have

$$H_{\nu+\frac{1}{2}} = \Big(I - \frac{h}{2}\Lambda_{\nu+\frac{1}{2}}\Big)^{-1} = \begin{pmatrix} 1+O(h) & 0 \\ 0 & O(\frac{\varepsilon}{h}) \end{pmatrix} \quad \text{for } \varepsilon \le Ch; \tag{4.11}$$

together with (4.4) this leads us to the desired estimate for the second term on the right hand side of (4.10):

$$h|Q_\nu|\Big\|H_{\nu+\frac{1}{2}}\Big(\bar{\jmath}_{\nu+1}-\bar{\jmath}_\nu\Big)\Big\| = O(h^{2q+3}) \tag{4.12}$$

(cf. (4.8) for the definition of $\bar{\jmath}_\nu$).

Now we estimate the first term on the right hand side of (4.10). By definition of $G_{\nu+\frac{1}{2}}$ and $H_{\nu\pm\frac{1}{2}}$ and due to (2.2),

$$\begin{aligned} G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} + Q_\nu H_{\nu+\frac{1}{2}} &= \Big(I - \frac{h}{2}\Lambda_{\nu+\frac{1}{2}}\Big)^{-1}\left[\Big(I + \frac{h}{2}\Lambda_{\nu+\frac{1}{2}}\Big)\Big(I - \frac{h}{2}\Lambda_{\nu-\frac{1}{2}}\Big)^{-1} + Q_\nu\, I\right] = \\ &= \begin{pmatrix} O(\frac{\varepsilon}{h}) + O(h) & 0 \\ 0 & O(\frac{\varepsilon}{h}) \end{pmatrix}. \end{aligned} \tag{4.13}$$

Therefore the first term on the right hand side of (4.10) can be estimated by

$$\begin{aligned} h\Big\|\Big(G_{\nu+\frac{1}{2}}H_{\nu-\frac{1}{2}} + Q_\nu H_{\nu+\frac{1}{2}}\Big)\bar{\jmath}_\nu\Big\||\Pi_{\nu-1}| &= \Big[\frac{\varepsilon}{h}\,O(h^{2q+2}) + O(h^{2q+3})\Big]|\Pi_{\nu-1}| = \\ &= \frac{\varepsilon}{h}\,O(h^{2q+2})|\Pi_{\nu-1}| + O(h^{2q+3}). \end{aligned} \tag{4.14}$$

Summarizing all that we obtain the recursion

$$\|\eta^+_{\nu+1}\| \le \Big(\frac{1+\omega}{1-\omega}\Big)^2\,\frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\,\|\eta^-_{\nu-1}\| + \frac{\varepsilon}{h}\,O(h^{2q+2})\,|\Pi_{\nu-1}| + O(h^{2q+3}) \tag{4.15}$$

with certain moderate O-constants. Estimating $\|\eta^-_{\nu-1}\|$ by $\|\eta^+_{\nu-1}\|$ on the basis of (3.17) and using once more (3.20) we end up with

$$\|\eta^+_{\nu+1}\| \le \underbrace{\Big(\frac{1+\omega}{1-\omega}\Big)^2\Big(\frac{1+\frac{h}{2}\vartheta}{1-\frac{h}{2}\vartheta}\Big)^2}_{=:\gamma^2=1+O(h)}\,\|\eta^+_{\nu-1}\| + \frac{\varepsilon}{h}\,O(h^{2q+2})\,|\Pi_{\nu-1}| + O(h^{2q+3}). \tag{4.16}$$

With

$$Q := \frac{1 - \frac{\kappa h}{2\varepsilon}}{1 + \frac{\kappa h}{2\varepsilon}} \tag{4.17}$$

(κ from (2.2)) we have $|\Pi_\nu| \le |Q|^\nu$. Thus, by induction,

$$\begin{aligned} \|\eta_\nu^+\| \;\le\; & \gamma^\nu \|\eta_0^+\| + O(h^{2q+2})\,\frac{\varepsilon}{h} \sum_{k=0(2)\nu-2} \gamma^{\nu-2-k}\, Q^k + \\ & + O(h^{2q+3}) \sum_{k=0(2)\nu-2} \gamma^k, \quad \nu = 2,\, 4,\, \ldots . \end{aligned} \tag{4.18}$$

Now the crucial point is that the accumulation of

$$\frac{\varepsilon}{h} \sum_{k=0(2)\nu-2} Q^k = \frac{\varepsilon}{h}\,\frac{1-Q^\nu}{1-Q^2} \tag{4.19}$$

does not generate a factor $1/h$ because

$$\frac{\varepsilon}{h}\,\frac{1}{1-Q^2} = \frac{\kappa}{8}\Big(1 + \frac{2\varepsilon}{\kappa h}\Big)^2 = O(1) \quad \text{uniformly for } \varepsilon \le Ch. \tag{4.20}$$

Therefore we end up with

$$\|\eta_\nu^+\| \le C_1^+(t_\nu)\|\eta_0^+\| + C_2^+(t_\nu) h^{2q+2} \tag{4.21}$$

with certain moderate bounds $C_i^+(t)$. Finally, the result (4.5) easily follows from (4.21) by definition of η_ν^+ (cf. (3.7)) and due to (2.25). □

Theorem 4.1. *Consider a stiff problem* (1.3) *of the type* (2.1)–(2.3) *and assume* $\varepsilon \le Ch$ *with* $C < \frac{\kappa}{2}$ *(*κ *from* (2.2)*). Denote*

$$\chi := \frac{\varepsilon}{h}. \tag{4.22}$$

Let $[0, t_E]$ *be a subinterval of the whole integration interval with constant stepsize* h *and assume inductively that - after transformation by* $T^{-1}(t)$ *- the accumulated global error* $T^{-1}(0)(\zeta_0 - z(0))$ *at* $t = 0$ *satisfies*

$$T^{-1}(0)\Big(\zeta_0 - z(0)\Big) = h^2 \begin{pmatrix} \tilde{x}_2(\chi^2) \\ \tilde{y}_2(\chi^2) \end{pmatrix} + O(h^4) \tag{4.23}$$

with certain moderate quantities $\tilde{x}_2(\chi^2)$ *and* $\tilde{y}_2(\chi^2)$ *which depend smoothly on* χ^2*. Then the global error of the implicit midpoint rule on* $[0, t_E]$ *can be written as follows:*

$$\zeta_\nu - z(t_\nu) = h^2 e_2(t_\nu; \chi^2) + R_\nu \tag{4.24}$$

where the function $e_2(t;\chi^2)$*, which depends smoothly on* χ^2*, is a smooth solution of the first variational equation in* (2.5)*. Furthermore,*

$$R_\nu = T(t_\nu) \begin{pmatrix} h^2 \varepsilon\, \check{\alpha}_2(t_\nu;\chi^2)\, \Pi_\nu \\ h^2 \Big[\beta_2(t_\nu;\chi^2) + \varepsilon\, \check{\beta}_2(t_\nu;\chi^2)\Big]\, \Pi_\nu \end{pmatrix} + O(h^4). \tag{4.25}$$

Here Π_ν is defined by (4.2)[10] *and $\breve{\alpha}_2(t;\chi^2)$, $\beta_2(t;\chi^2)$ and $\breve{\beta}_2(t;\chi^2)$ are smooth functions in t which depend smoothly on χ^2. $\beta_2(t;\chi^2)$ is the solution of*

$$\begin{aligned} \beta_2'(t;\chi^2) &= d_{22}^{(1)}(t)\,\beta_2(t;\chi^2), \\ \beta_2(0;\chi^2) &= \tilde{y}_2(\chi^2) - Y_{0,2}(0) \end{aligned} \tag{4.26}$$

(with $d_{22}^{(1)}(t)$ from (3.32) and where $Y_{0,2}(0)$ is the starting value of the leading term of the ε-expansion of the second component of $\bar{e}_2(t) = T^{-1}(t)e_2(t)$). The equations defining $\breve{\alpha}_2(t;\chi^2)$ and $\breve{\beta}_2(t;\chi^2)$ are given in (4.40) and (4.42) below.

Proof. We consider the "expanded" transformed remainder equation (2.38) for $\rho_\nu = \begin{pmatrix} x_\nu \\ y_\nu \end{pmatrix}$ and use the same scaling as in the proof of Theorem 3.1 (cf. (3.33), (3.34)). It is now convenient to rewrite the second component (3.34) as

$$\begin{aligned}
&\frac{1}{h}\Big[\; -\frac{h^2}{4}d_{21}^{(2)}(\hat{t}_\nu) - \frac{h^4}{96}d_{21}^{(4)}(\hat{t}_\nu) + \ldots\Big]\Big[\frac{\varepsilon}{h}\Big(x_\nu - x_{\nu-1}\Big) + c_2(\hat{t}_\nu)\,\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big)\Big] + \\
&+\frac{1}{h}\Big[1 - \frac{h^2}{4}d_{22}^{(2)}(\hat{t}_\nu) - \frac{h^4}{96}d_{22}^{(4)}(\hat{t}_\nu) + \ldots\Big]\Big[\frac{\varepsilon}{h}\Big(y_\nu - y_{\nu-1}\Big) + c_2(\hat{t}_\nu)\,\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big)\Big] = \\
= \;&\Big[d_{21}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{21}^{(3)}(\hat{t}_\nu) + \ldots\Big]\Big[\frac{\varepsilon}{h}\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big) + \frac{c_2(\hat{t}_\nu)}{4}\Big(x_\nu - x_{\nu-1}\Big)\Big] + \\
&+\Big[d_{22}^{(1)}(\hat{t}_\nu) + \frac{h^2}{8}d_{22}^{(3)}(\hat{t}_\nu) + \ldots\Big]\Big[\frac{\varepsilon}{h}\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big) + \frac{c_2(\hat{t}_\nu)}{4}\Big(y_\nu - y_{\nu-1}\Big)\Big] + \\
&+h^2\,\frac{\varepsilon}{h}\,\gamma(\hat{t}_\nu)\,\frac{1}{2}\Big(x_{\nu-1} + x_\nu\Big) + h^2\,\frac{\varepsilon}{h}\,\gamma(\hat{t}_\nu)\,\frac{1}{2}\Big(y_{\nu-1} + y_\nu\Big) + \ldots - \\
&-\frac{h^3}{4}\,\frac{\varepsilon}{h}\,\gamma(\hat{t}_\nu)\Big(x_\nu - x_{\nu-1}\Big) - \frac{h^3}{4}\,\frac{\varepsilon}{h}\,\gamma(\hat{t}_\nu)\Big(y_\nu - y_{\nu-1}\Big) + \ldots + \\
&+\frac{\varepsilon}{h}\cdot(\textit{higher order nonlinear terms}) - \frac{1}{h}c_2(\hat{t}_\nu)\bar{a}_{\nu,2} + \frac{\varepsilon}{h}\bar{b}_{\nu,2} - \frac{\varepsilon}{h}\bar{c}_{\nu,2}
\end{aligned} \tag{4.27}$$

In the following any occurrence of ε is rewritten as $\varepsilon = h\chi$ where $\chi \le C$ is moderate under the present assumption $\varepsilon \le Ch$. (Recall that, in contrast to section 3, the parameter χ is now defined as $\chi := \varepsilon/h$ - cf. (4.22).)

The starting condition to be fulfilled is (cf. (4.24))

$$\rho_0 = T^{-1}(0)\Big(\zeta_0 - z(0)\Big) - h^2\bar{e}_2(0;\chi^2), \tag{4.28}$$

thus, due to our inductive assumption (4.23) and due to (2.24):

$$\begin{aligned}
\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} &= h^2\begin{pmatrix} \tilde{x}_2(\chi^2) - X_{0,2}(0;\chi^2) - \varepsilon X_{1,2}(0;\chi^2) - \ldots \\ \tilde{y}_2(\chi^2) - Y_{0,2}(0) - \varepsilon Y_{1,2}(0;\chi^2) - \ldots \end{pmatrix} + O(h^4) = \\
&= h^2\begin{pmatrix} \tilde{x}_2(\chi^2) - X_{0,2}(0;\chi^2) \\ \tilde{y}_2(\chi^2) - Y_{0,2}(0) \end{pmatrix} + h^3\begin{pmatrix} -\chi X_{1,2}(0;\chi^2) \\ -\chi Y_{1,2}(0;\chi^2) \end{pmatrix} + O(h^4).
\end{aligned} \tag{4.29}$$

[10] Note that $\Pi_\nu \neq 0$ by assumption $\varepsilon \le Ch < \frac{\kappa}{2}h$.

The quantities $X_{\ell,2}(0;\chi^2)$ and $Y_{\ell,2}(0;\chi^2)$ are still undetermined. In (4.36) and (4.41) below we shall make a special choice for the $X_{\ell,2}(0;\chi^2)$. By construction of smooth solutions (cf. section 2) the $Y_{\ell,2}(0;\chi^2)$ are then fixed.

For x_ν and y_ν we make the ansatz

$$\begin{aligned}\begin{pmatrix} x_\nu \\ y_\nu \end{pmatrix} &= \begin{pmatrix} h^2\varepsilon\,\breve{\alpha}_2(t_\nu;\chi^2)\,\Pi_\nu + x_\nu^{(rem)} \\ h^2\Big[\beta_2(t_\nu;\chi^2) + \varepsilon\,\breve{\beta}_2(t_\nu;\chi^2)\Big]\,\Pi_\nu + y_\nu^{(rem)} \end{pmatrix} = \\ &= \begin{pmatrix} h^3\chi\,\breve{\alpha}_2(t_\nu;\chi^2)\,\Pi_\nu + x_\nu^{(rem)} \\ h^2\,\beta_2(t_\nu;\chi^2)\,\Pi_\nu + h^3\chi\,\breve{\beta}_2(t_\nu;\chi^2)\,\Pi_\nu + y_\nu^{(rem)} \end{pmatrix}\end{aligned} \tag{4.30}$$

with certain functions $\breve{\alpha}_2(t;\chi^2)$, $\beta_2(t;\chi^2)$ and $\breve{\beta}_2(t;\chi^2)$ which are to be determined. Taylor expansion about $\hat{t}_\nu$ and using the definition of Π_ν (cf. (4.2)) yields

$$\begin{aligned}&\beta(t_\nu;\chi^2)\,\Pi_\nu - \beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1} = \\ &= \left[2\beta(\hat{t}_\nu;\chi^2) - h\frac{2\chi}{c_2(\hat{t}_\nu)}\beta'(\hat{t}_\nu;\chi^2) + O(h^2)\right]\cdot\left(-\Pi_{\nu-1}\frac{1}{1+\frac{2\chi}{c_2(\hat{t}_\nu)}}\right),\end{aligned} \tag{4.31}$$

$$\begin{aligned}&\frac{1}{2}\Big(\beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1} + \beta(t_\nu;\chi^2)\,\Pi_\nu\Big) = \\ &= \left[-\frac{2\chi}{c_2(\hat{t}_\nu)}\beta(\hat{t}_\nu;\chi^2) + \frac{h}{2}\,\beta'(\hat{t}_\nu;\chi^2) + O(h^2)\right]\cdot\left(-\Pi_{\nu-1}\frac{1}{1+\frac{2\chi}{c_2(\hat{t}_\nu)}}\right)\end{aligned} \tag{4.32}$$

and

$$\begin{aligned}&\chi\Big(\beta(t_\nu;\chi^2)\,\Pi_\nu - \beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1}\Big) + c_2(\hat{t}_\nu)\,\frac{1}{2}\Big(\beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1} + \beta(t_\nu;\chi^2)\,\Pi_\nu\Big) = \\ &= \Big[h\,\beta'(\hat{t}_\nu;\chi^2) + O(h^3)\Big]\cdot\Big(\chi - \frac{c_2(\hat{t}_\nu)}{2}\Big)\Pi_{\nu-1},\end{aligned} \tag{4.33}$$

$$\begin{aligned}&\chi\,\frac{1}{2}\Big(\beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1} + \beta(t_\nu;\chi^2)\,\Pi_\nu\Big) + \frac{c_2(\hat{t}_\nu)}{4}\,\Big(\beta(t_\nu;\chi^2)\,\Pi_\nu - \beta(t_{\nu-1};\chi^2)\,\Pi_{\nu-1}\Big) = \\ &= \Big[\beta(\hat{t}_\nu;\chi^2) + O(h^2)\Big]\cdot\Big(\chi - \frac{c_2(\hat{t}_\nu)}{2}\Big)\Pi_{\nu-1}\end{aligned} \tag{4.34}$$

(analogously for α). After this preparation we insert into (3.33) and (4.27) and equate coefficients of h^2 and h^3 (omitting certain common nonvanishing factors).

- *Coefficients of* h^2:
  There is only an equation for β_2:

  $$\beta_2'(\hat{t}_\nu;\chi^2) = d_{22}^{(1)}(\hat{t}_\nu)\,\beta_2(\hat{t}_\nu;\chi^2). \tag{4.35}$$

  To satisfy the starting condition (4.29) at the h^2-level, the following choice has to be made:

  - In the first component we set

    $$X_{0,2}(0;\chi^2) := \tilde{x}_2(\chi^2) \tag{4.36}$$

    (recall that for a smooth solution $e_2(t)$ the $X_{\ell,2}(0)$ can be chosen arbitrarily - cf. section 2).

– In the second component the $Y_{\ell,2}(0;\chi^2)$ are fixed by the requirement that $e_2(t;\chi^2)$ be smooth (cf. section 2). (In particular, $Y_{0,2}(0)$ is defined in terms of data functions - cf. (2.22) - and is therefore χ-independent.) So, due to (4.29), $\beta_2(0;\chi^2)$ is fixed by $\beta_2(0;\chi^2) = \tilde{y}_2(\chi^2) - Y_{0,2}(0)$.

Thus we see from (4.35) that $\beta_2(t;\chi^2)$ is the solution of the initial value problem

$$\begin{aligned} \beta_2'(t;\chi^2) &= d_{22}^{(1)}(t)\,\beta_2(t;\chi^2), \\ \beta_2(0;\chi^2) &= \tilde{y}_2(\chi^2) - Y_{0,2}(0) \end{aligned} \tag{4.37}$$

which is smooth under the present assumption $\varepsilon \leq Ch$ ($\chi \leq C$). Note that - since the coefficient function $d_{22}^{(1)}(t)$ in (4.37) is independent of χ and, by inductive assumption, the starting value for β_2 depends smoothly on χ^2 - β_2 itself depends smoothly on the parameter χ^2.

- *Coefficients of* h^3:

$$2\chi\,\breve{\alpha}_2(\hat{t}_\nu;\chi^2) + 2\chi\,\frac{d_{12}^{(1)}(\hat{t}_\nu)}{c_2(\hat{t}_\nu)}\,\beta_2(\hat{t}_\nu;\chi^2) = 0, \tag{4.38}$$

$$\chi\,\breve{\beta}_2'(\hat{t}_\nu;\chi^2) = \chi\,d_{21}^{(1)}(\hat{t}_\nu)\,\breve{\alpha}_2(\hat{t}_\nu;\chi^2) + \chi\,d_{22}^{(1)}(\hat{t}_\nu)\,\breve{\beta}_2(\hat{t}_\nu;\chi^2). \tag{4.39}$$

(4.38) is an *algebraic* equation fixing $\breve{\alpha}_2(t;\chi^2)$ (which obviously depends smoothly on χ^2):

$$\breve{\alpha}_2(t;\chi^2) = -\frac{d_{12}^{(1)}(t)}{c_2(t)}\,\beta_2(t;\chi^2). \tag{4.40}$$

Thus also $\breve{\alpha}_2(0;\chi^2)$ is fixed; to satisfy the starting condition (4.29) at the h^3-level in the first component we make use of the freedom of choosing $X_{1,2}(0;\chi^2)$ (recall again that a smooth solution $e_2(t)$ exists for an arbitrary choice of the $X_{\ell,2}(0)$):

$$X_{1,2}(0;\chi^2) := -\breve{\alpha}_2(0;\chi^2) = \frac{d_{12}^{(1)}(0)}{c_2(0)}\,\beta_2(0;\chi^2). \tag{4.41}$$

The starting value for the second component is fixed by (4.29): $\breve{\beta}_2(0;\chi^2) = -Y_{1,2}(0;\chi^2)$. Thus we see from (4.39) and (4.40) that $\breve{\beta}_2(t;\chi^2)$ is the solution of the smooth initial value problem

$$\begin{aligned} \breve{\beta}_2'(t;\chi^2) &= d_{22}^{(1)}(t)\,\breve{\beta}_2(t;\chi^2) - \frac{d_{12}^{(1)}(t)\,d_{21}^{(1)}(t)}{c_2(t)}\,\beta_2(t;\chi^2), \\ \breve{\beta}_2(0;\chi^2) &= -Y_{1,2}(0;\chi^2). \end{aligned} \tag{4.42}$$

It is obvious that $\breve{\beta}_2(t;\chi^2)$ depends smoothly on the parameter χ^2.

- *Estimation of the* h^4*-remainder term:*
To estimate the remainder term

$$\rho_\nu^{(rem)} = \begin{pmatrix} x_\nu^{(rem)} \\ y_\nu^{(rem)} \end{pmatrix} \tag{4.43}$$

one can proceed in a completely analogous way as in the proof of Theorem 3.1 (cf. (3.49) ff.). Due to the re-scaling of (3.33) and (4.27) by $1/h$ and h/ε, resp., the inhomogeneity of the difference equation for $\rho_\nu^{(rem)}$ is easily seen to be of the following form:

$$\Lambda_{\nu-\frac{1}{2}}\bar{a}_\nu + \bar{b}_\nu - \bar{c}_\nu + h^4\delta_\nu + \begin{pmatrix} h^3\sigma_\nu\Pi_{\nu-1} \\ \frac{h^5}{\varepsilon}\tau_\nu\Pi_{\nu-1} \end{pmatrix} + O(h^4). \tag{4.44}$$

Here, $h^4\delta_\nu + \begin{pmatrix} h^3\sigma_\nu\Pi_{\nu-1} \\ \frac{h^5}{\varepsilon}\tau_\nu\Pi_{\nu-1} \end{pmatrix} + O(h^4)$ originates from the collection of all terms involving $\breve{\alpha}_2(\hat{t}_\nu;\chi^2)$, $\beta_2(\hat{t}_\nu;\chi^2)$ and $\breve{\beta}_2(\hat{t}_\nu;\chi^2)$ which have not yet appeared in (4.35), (4.38) and (4.39) above. Note that due to (4.31)–(4.34) all these terms contain indeed a factor $\Pi_{\nu-1}$; the factors $h^3 = h^4 \cdot \frac{1}{h}$ and $h^5/\varepsilon = h^4 \cdot \frac{h}{\varepsilon}$, resp., originate from the re-scaling of (3.33) and (4.27). σ_ν and τ_ν are certain moderate, χ-dependent quantities, σ_ν satisfying the smoothness assumption (4.4).
Since $\rho_0^{(rem)} = O(h^4)$ by construction, the desired estimate $\rho_\nu^{(rem)} = O(h^4)$ follows now by application of Lemma 4.1. □

We have shown that the behavior of the nonsmooth terms of the global error is characterized by the factor Π_ν. If even $\varepsilon \leq Ch^2$ holds, Π_ν is oscillating with a smoothly varying amplitude. For $Ch^2 \leq \varepsilon \leq \frac{\kappa}{2}h$, Π_ν is still oscillating but with a nonsmooth, significantly *decreasing* amplitude. This damping property becomes stronger as ε/h increases.

Concerning the inductive application of Theorem 4.1 w.r.t. several subintervals with constant stepsizes, analogous remarks apply as given after the proof of Theorem 3.1 (section 3). To justify the inductive assumption (4.23) it has to be verified that for fixed $t \equiv t_\nu$, ν even, Π_ν depends smoothly on χ^2; this is shown in (4.45)–(4.47) below. Thus, the starting value for the next subinterval depends indeed smoothly on χ^2, as assumed in (4.23). In cases where the subintervals are sufficiently long - such that at a point where the stepsize changes the terms involving Π_ν are damped out to the $O(h^4)$-level - the inductive starting values are actually not influenced by χ^2; therefore a look at the proof of Theorem 4.1 shows that $e_2(t)$ is χ-independent in this case.
To see that Π_ν depends indeed smoothly on $\chi^2 = (\frac{\varepsilon}{h})^2$, we rewrite Π_ν as

$$\Pi_\nu = \prod_{\ell=1}^{\nu} \left(\frac{1 - \frac{2\chi}{c_2(\hat{t}_\ell)}}{1 + \frac{2\chi}{c_2(\hat{t}_\ell)}} \right) \cdot \underbrace{(-1)^\nu}_{1} = \exp\left(h \sum_{\ell=1}^{t/h} \frac{1}{h} \ln\left(\frac{1 - \frac{2\chi}{c_2(\hat{t}_\ell)}}{1 + \frac{2\chi}{c_2(\hat{t}_\ell)}} \right) \right). \tag{4.45}$$

The sum occurring in (4.45) is the midpoint quadrature for

$$\frac{1}{\varepsilon} \int_0^t \chi \ln\left(\frac{1 - \frac{2\chi}{c_2(s)}}{1 + \frac{2\chi}{c_2(s)}} \right) ds. \tag{4.46}$$

(Note that the factor $1/\varepsilon$ is not critical because the integrand is negative and so $\exp(\frac{1}{\varepsilon}\int_0^t \ldots ds)$ is moderate.) For each s the integrand of (4.46) is an even function in χ:

$$\chi \ln\left(\frac{1 - \frac{2\chi}{c_2(s)}}{1 + \frac{2\chi}{c_2(s)}} \right) = -c_2(s)\, \frac{2\chi}{c_2(s)} \left[\frac{2\chi}{c_2(s)} + \frac{1}{3}\left(\frac{2\chi}{c_2(s)}\right)^3 + \frac{1}{5}\left(\frac{2\chi}{c_2(s)}\right)^5 + \ldots \right], \tag{4.47}$$

and the same is true for $\int_0^t \ldots ds$. The quadrature error can be expanded in powers of h^2 in a well-known way; the coefficients of this expansion involve odd derivatives of the integrand at $s = 0$ and $s = t$ which again depend smoothly on χ^2.

Concerning *longer expansions* than covered by Theorem 4.1 the argumentation is the same as in the above proof. Due to nonlinear effects, however, not only terms with the factor Π_ν appear but also higher powers Π_ν^ℓ, $\ell > 1$. In contrast to section 3 (where $(-1)^\nu$ appeared instead of Π_ν), even powers of Π_ν (which do not oscillate) are not smooth but also significantly decaying (unless even $\varepsilon \leq Ch^2$ holds). To describe the error structure in the nonlinear case it is necessary to introduce a scale of terms

$$\textit{smooth function}\,(t_\nu) \cdot \Pi_\nu^\ell, \quad \ell = 1, 2, \ldots \tag{4.48}$$

within the ansatz for R_ν. Obviously, the error structure is less regular as for $\varepsilon \leq Ch^2$; however, with decrasing h (and in particular for $h \to 0$ which is not discussed in detail here) the damping properties of the Π_ν^ℓ become stronger such that sufficiently away from $t = 0$ the global error is essentially given by $h^2 e_2(t_\nu) + h^4 e_4(t_\nu) + \ldots$.

References

[1] W. Auzinger, R. Frank, F. Macsek, *Asymptotic error expansions for stiff equations: The implicit Euler scheme*, SIAM J. Numer. Anal. 27, 67–104 (1990).

[2] W. Auzinger, R. Frank, *Asymptotic error expansions for stiff equations: An analysis for the implicit midpoint and trapezoidal rules in the strongly stiff case*, Numer. Math. 56, 469–499 (1989).

[3] W. Auzinger, R. Frank, G. Kirlinger, *Asymptotic error expansions for stiff equations: Applications*, Computing 43, 223–253 (1990).

[4] W. Auzinger, R. Frank, G. Kirlinger, *A note on convergence concepts for stiff problems*, Report Nr. 80/89, Institut fuer Angewandte und Numerische Mathematik, TU Wien, 1989 (to appear in Computing 44, 1990).

[5] G. Bader, P. Deuflhard, *A semi-implicit mid-point rule for stiff systems of ordinary differential equations*, Numer. Math. 41, 373–398 (1983).

[6] G. Dahlquist, B. Lindberg, *On some implicit one-step methods for stiff differential equations*, Dept. of Computer Sciences, Royal Institute of Technology, Report TRITA-NA-7302, 1973.

[7] K. Dekker, J. G. Verwer, *Stability of Runge-Kutta methods for stiff nonlinear differential equations*, North-Holland, Amsterdam, New York, 1984.

[8] W. B. Gragg, *Repeated extrapolation to the limit in the numerical solution of ordinary differential equations*, Thesis, UCLA 1963.

[9] E. Hairer, Ch. Lubich, *Extrapolation at stiff differential equations*, Numer. Math. 52, 377–400 (1988).

[10] J. F. B. M. Kraaijevanger, *B-Convergence of the implicit midpoint rule and the trapezoidal rule*, BIT 25, 652–666 (1985).

[11] A. H. Nayfeh, D. T. Mook, *Nonlinear oscillations*, J. Wiley & Sons, New York, 1979.

[12] R. E. O'Malley Jr., *Introduction to singular perturbations*, Academic Press, New York - London - Toronto - Sydney - San Francisco, 1974.

[13] D. R. Smith, *Singular perturbation theory*, Cambridge University Press, Cambridge - London - New York, 1985.

[14] H. J. Stetter, *Asymptotic expansions for the error of discretization algorithms for non-linear functional equations*, Numer. Math. 7, 18–31 (1965).

[15] M. Van Veldhuizen, *Asymptotic expansions of the global error for the implicit midpoint rule (stiff case)*, Computing 33, 185–192 (1984).