Zuse Institute Berlin

Takustr. 7
14195 Berlin
Germany

ANDREAS GRIEWANK
RICHARD HASENFELDER
MANUEL RADONS
LUTZ LEHMANN
TOM STREUBEL

Integrating Lipschitzian Dynamical Systems using Piecewise Algorithmic Differentiation

ZIB Report 17-44 (July 2017)

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30-84185-0
Telefax: +49 30-84185-125

E-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Integrating Lipschitzian Dynamical Systems using Piecewise Algorithmic Differentiation

Andreas Griewank[1], Richard Hasenfelder[2], Manuel Radons[3], Lutz Lehmann[2], and Tom Streubel[4,2]

[1]School of Mathematical Sciences and Information Technology, Ecuador
[2]Humboldt University of Berlin, Germany
[3]Technical University in Berlin, Germany
[4]Zuse Institute Berlin, Germany

corresp. author: hasenfel@math.hu-berlin.de

Abstract

In this article we analyze a generalized trapezoidal rule for initial value problems with piecewise smooth right hand side $F : \mathbb{R}^n \to \mathbb{R}^n$ based on a generalization of algorithmic differentiation. When applied to such a problem, the classical trapezoidal rule suffers from a loss of accuracy if the solution trajectory intersects a nondifferentiability of F. The advantage of the proposed generalized trapezoidal rule is threefold: Firstly, we can achieve a higher convergence order than with the classical method. Moreover, the method is energy preserving for piecewise linear Hamiltonian systems. Finally, in analogy to the classical case we derive a third order interpolation polynomial for the numerical trajectory. In the smooth case the generalized rule reduces to the classical one. Hence, it is a proper extension of the classical theory. An error estimator is given and numerical results are presented.

Keywords Automatic Differentiation, Lipschitz Continuity, Piecewise Linearization, Nonsmooth, Trapezoidal Rule, Energy Preservation, Dense Output
MSC 2010 65L05, 65L06, 65L70, 65L99, 65P10

1 Introduction

Many realistic computer models are nondifferentiable in that the functional relation between input and output variables is not smooth. We are particularly focusing on Lipschitz continuous models where the nondifferentiabilities have a special structure. We present a technique that handles such functions in the context of the numerical solution of ordinary differential equations (ODE's). It is based on algorithmic differentiation (AD) and generalizes this concept. For further information on the general theory of AD we refer to [GW08, Nau12].
Consider the following initial value problem for an autonomous ODE.

$$\dot{x}(t) = F(x(t)), \quad x(0) = x_0 \ , \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be locally Lipschitz continuous. It is well known that this system has a unique local solution up to some time $\bar{t} > 0$. For a time-step $h > 0$ the exact solution of (1) satisfies

$$\hat{x} = \check{x} + \int_0^h F(x(t))\mathrm{d}t \ ,$$

with $\hat{x} = x(h)$ and $\check{x} = x_0$. In general the integral cannot be evaluated exactly.

In the derivation of the classical trapezoidal rule a linear approximation of the right hand side is utilized. The integration of these approximations yields a third order local truncation error if F is smooth. If F is only Lipschitz continuous, the truncation error will drop to second order where the solution trajectory intersects a nondifferentiability.

The key idea to reestablish a third order truncation error everywhere is to approximate F by a **piecewise linear**[1] function that reflects the structure of the nondifferentiabilities of F. Employing this approach we will construct a generalized trapezoidal rule with the following three major benefits:

- We achieve second order global accuracy in general and third order via Romberg extrapolation along solution trajectories with finitely many kink locations.

[1] Our notion of linearity includes nonhomogeneous functions, where the adjective *affine* or perhaps *polyhedral* would be more precise. However, such mathematical terminology might be less appealing to computational practitioners and to the best of our knowledge there are no good nouns corresponding to *linearity* and *linearization* for the adjectives affine and polyhedral.

- A third order interpolating polynomial as a continuous approximation of the trajectory will be given.
- The method is energy preserving on piecewise linear Hamiltonian systems.

Content and Structure

The article is organized as follows: In Section 2 we introduce the necessary prerequisites from piecewise linear theory and generalized algorithmic differentiation. In Section 3 the generalized trapezoidal rule is constructed and convergence results are proved. The extrapolation results are presented in Section 4, as well as the geometric integration properties. The error estimator is derived in Section 5. The sixth section contains numerical results. We conclude with some final remarks.

2 Piecewise Linear Model

Definition 2.1. *A continuous function $F : \mathbb{R}^n \to \mathbb{R}^m$ is called* **piecewise linear** *if there exists a finite number of affine* **selection functions** $F_i : \mathbb{R}^n \to \mathbb{R}^m$ *such that at any given $x \in \mathbb{R}^n$ there exists at least one index i with $F(x) = F_i(x)$.*

Let the index set $I = \{1, ..., k\}$ of the selection functions be given. According to [Sch12, Prop. 2.2.2] we can find subsets $M_1, ..., M_l \subset I$ such that a scalar valued piecewise linear function f can be represented as

$$f(x) = \max_{1 \le i \le l} \min_{j \in M_i} f_j(x) \, .$$

This concept, which is called **max-min representation**, naturally carries over to vector valued functions F, where we can find such a decomposition for every component of the image. The special type of piecewise linear functions that will be utilized here will naturally have this representation.

Note that piecewise linear functions are globally Lipschitz continuous. For a further discussion of their properties we refer to [Sch12]. Next we consider continuous, piecewise differentiable functions F that can be computed by a finite program called an *evaluation procedure*. An evaluation procedure is a composition of so-called elementary functions which make up the atomic constituents of more complex functions. Basically, the selection of elementary functions for the library is arbitrary, as long as they comply with assumption (ED) (elementary differentiability, in [GW08]), meaning that

they are at least once Lipschitz-continuously differentiable on their valid open domains. Common examples are:

$$\tilde{\Phi} \ := \ \{+, -, *, /, \sin, \cos, \tan, \cot, \exp, \log, \dots\} \ .$$

In our case, we will allow the evaluation procedure of $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ to contain, in addition to the usual smooth elementary functions, the absolute value $\mathrm{abs}(x) = |x|$, i.e., our library is of the form

$$\Phi \ := \tilde{\Phi} \ \cup \ \{\mathrm{abs}\} \ .$$

Consequently, we can also handle the maximum and minimum of two values via the representation

$$\max(u, v) = (u + v + |u - v|)/2, \quad \min(u, v) = (u + v - |u - v|)/2 \ .$$

We call the resulting functions *composite piecewise differentiable*. These functions are locally Lipschitz continuous and almost everywhere differentiable in the classical sense. Furthermore, they are differentiable in the sense of Bouligand and Clarke, cf. [Cla83]. The evaluation procedure of $y = F(x)$ can be interpreted as a directed, acyclic graph from $x = (v_{1-n}, ..., v_0)$ to $y = (v_{l-m+1}, ..., v_l)$, where the intermediate values $v_i, i = 1, ..., l$ are computed by binary operations $v_i = v_j \circ v_k$ with $\circ \in \{+, -, *\}$ and $v_j, v_k \prec v_i$ or unary functions $v_i = \phi_i(v_j)$ with $v_j \prec v_i$, where $\phi_i \in \Phi$. The relation $\prec$ represents the data dependence in the graph of the evaluation procedure, which must be acyclic.

We now want to compute an incremental approximation $\Delta y = \Delta F(\mathring{x}; \Delta x)$ to $F(\mathring{x} + \Delta x) - F(\mathring{x})$ at a given $\mathring{x}$ and for a variable increment Δx. Assuming that all functions other than the absolute value are differentiable, we introduce the propagation rules

$$\begin{array}{ll} \Delta v_i = \Delta v_j \pm \Delta v_k & \text{for } \mathring{v}_i = \mathring{v}_j \pm \mathring{v}_k \ , \\ \Delta v_i = \mathring{v}_j * \Delta v_k + \Delta v_j * \mathring{v}_k & \text{for } \mathring{v}_i = \mathring{v}_j * \mathring{v}_k \ , \\ \Delta v_i = \mathring{c}_{ij} \Delta v_j \quad \text{with } \mathring{c}_{ij} = \varphi'(\mathring{v}_j) & \text{for } \mathring{v}_i = \varphi_i(\mathring{v}_j) \neq \mathrm{abs}(\cdot) \ , \\ \Delta v_i = \mathrm{abs}(\mathring{v}_j + \Delta v_j) - \mathrm{abs}(\mathring{v}_j) & \text{for } \mathring{v}_i = \mathrm{abs}(\mathring{v}_j) \ . \end{array} \tag{2}$$

Whenever F is globally differentiable (i.e., there are no abs calls in the evaluating procedure) we get $\Delta y = F'(\mathring{x})\Delta x$, where $F'(\mathring{x}) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix.

Note that the propagation rules (2) rely on the so-called tangent approximation of F at a certain point $\mathring{x}$. However, there are applications of piecewise linearization (especially concerning ODE integration) where one

wants to consider approximations of F based on secants. Given two points $\check{x}, \hat{x}$ we compute $\mathring{x} = (\check{x} + \hat{x})/2$ and $\mathring{F} = (F(\check{x}) + F(\hat{x}))/2$. Now we consider the secant approximation of F:

$$F(x) \approx \mathring{F} + \Delta F(\check{x}, \hat{x}; x - \mathring{x}) \; . \tag{3}$$

The two piecewise linearization modes are displayed in Figure 1 a) and Figure 1 b), respectively. The figures show a simple example, where $F \colon \mathbb{R} \to \mathbb{R}$ is the maximum of two smooth functions. In this specific case the tangent approximation is simply the maximum of the two tangents at $\mathring{x}$. Similarly, the secant approximation is the maximum of the two secants through $F_i(\check{x})$ and $F_i(\hat{x})$ for the selection functions F_1 and F_2, respectively.
In order to utilize AD for the algorithmic computation of the secant approximation in (3) we observe that in (2) the intermediate values can be regarded as functions evaluated at the unique reference point $\mathring{x}$, with $\mathring{v}_i = v_i(\mathring{x})$. Now consider as this reference point the midpoint $\mathring{x} = (\check{x} + \hat{x})/2$ such that the intermediate values are

$$\mathring{v}_i = (\check{v}_i + \hat{v}_i)/2 \; , \text{ with } \; , \check{v}_i = v_i(\check{x}), \hat{v}_i = v_i(\hat{x}) \; .$$

Replacing $\mathring{v}_i$ in (2) with this expression based on $\check{x}$ and $\hat{x}$, we observe that the first and second line are the same for the secant linearization. The third line has to be changed slightly, since the tangent slope $\mathring{c}_{ij}$ has to be replaced by the secant slope

$$\mathring{c}_{ij} = \begin{cases} \phi_i'(\mathring{v}_j) & \text{if } \check{v}_j = \hat{v}_j \\ \dfrac{\hat{v}_i - \check{v}_i}{\hat{v}_j - \check{v}_j} & \text{otherwise} \end{cases} \; .$$

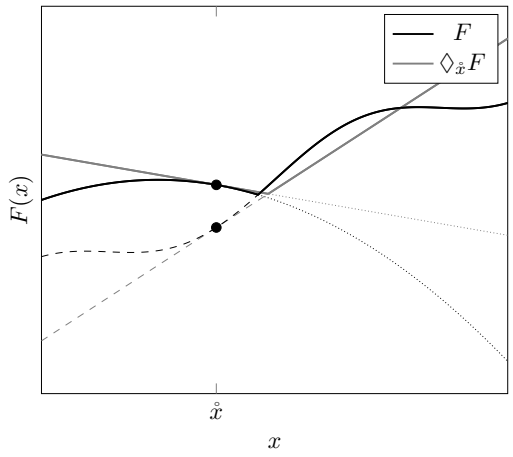
The last rule is left unchanged except that now $\mathring{v}_i = (\check{v}_i + \hat{v}_i)/2 = (|\check{v}_j| + |\hat{v}_j|)/2$. Note that, if $\check{x} = \hat{x}$, we obtain $\Delta F(\mathring{x}, \Delta x) = \Delta F(\check{x}, \hat{x}; \Delta x)$. A complete discussion on this implementation topic can be found in [Gri13, Sec. 7]. Additionally, a division-free implementation and thus numerically stable implementation is discussed in [GSHR, Section 6].

In contrast to the presentation in our previous papers we will now also use the nonincremental forms of the tangent approximation

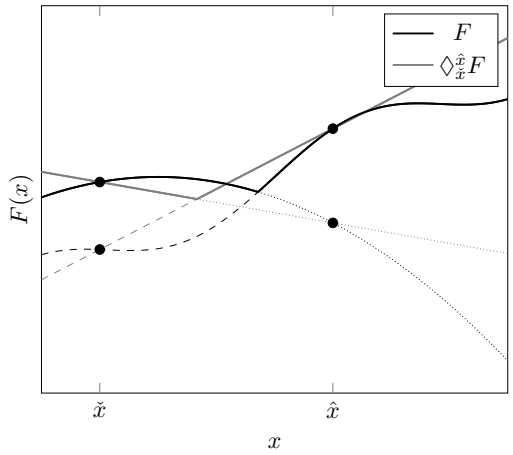
$$\Diamond_{\mathring{x}} F(x) \; \equiv \; F(\mathring{x}) + \Delta F(\mathring{x}; x - \mathring{x}) \tag{4}$$

and the secant approximation

$$\Diamond_{\check{x}}^{\hat{x}} F(x) \; \equiv \; \tfrac{1}{2}(F(\check{x}) + F(\hat{x})) + \Delta F(\check{x}, \hat{x}; x - \mathring{x}) \; , \tag{5}$$



(a) Tangent mode linearization



(b) Secant mode linearization

Figure 1: Piecewise linearization modes

respectively.

Hereafter, we will denote by $\|\cdot\| \equiv \|\cdot\|_\infty$ the infinity norm. Due to norm equivalence in finite dimensional spaces all inequalities to be derived take the same form in other norms, provided the constants are adjusted accordingly.

Moreover, we will frequently use the following central Proposition 4.2 from [GSHR]:

Proposition 2.1 (Griewank et.al, Prop. 4.2)**.** *Suppose $x, \check{x}, \hat{x}, \tilde{x}, \check{y}, \hat{y}, \check{z}, \hat{z} \in \mathbb{R}^n$ are restricted to a sufficiently small closed convex neighborhood $K \subseteq \mathbb{R}^n$ where the evaluation procedure for $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is well defined. Then there are Lipschitz constants β_F and γ_F, such that*

(i) Lipschitz continuity of function and piecewise linear models

$$\max\left(\|F(x) - F(\tilde{x})\|, \|\Diamond_{\check{x}}^{\hat{x}} F(x) - \Diamond_{\check{x}}^{\hat{x}} F(\tilde{x})\|, \|\Diamond_{\mathring{x}} F(x) - \Diamond_{\mathring{x}} F(\tilde{x})\|\right) \;\le\; \beta_F \|x - \tilde{x}\|$$

The constant β_F can be defined by the recurrences $\beta_v = \beta_u + \beta_w$ if $v = u + v$, $\beta_v = \beta_u$ if $v = |u|$ and

$$\beta_v = \beta_u L_K(\varphi) \quad \text{if} \quad v = \varphi(u) \quad \text{with} \quad L_K(\varphi) \equiv \max_{x \in K} |\varphi'(u(x))|$$

(ii) Error between function and secant or tangent model

$$\begin{aligned} \|F(x) - \Diamond_{\check{x}}^{\hat{x}} F(x)\| &\le \tfrac{1}{2}\gamma_F \|x - \hat{x}\| \|x - \check{x}\| \\ \|F(x) - \Diamond_{\mathring{x}} F(x)\| &\le \tfrac{1}{2}\gamma_F \|x - \mathring{x}\|^2 \end{aligned}$$

The constant γ_F can be defined using the recurrences $\gamma_v = \gamma_u + \gamma_w$ if $v = u + v$, $\gamma_v = \gamma_u$ if $v = |u|$ and

$$\gamma_v = L_K(\varphi)\gamma_u + L_K(\varphi')\beta_u^2 \quad \text{if} \quad v = \varphi(u) \quad \text{with} \quad L_K(\varphi') \equiv \max_{x \in K} |\varphi''(u(x))|$$

(iii) Lipschitz continuity of secant and tangent model

$$\begin{aligned} \|\Diamond_{\check{z}}^{\hat{z}} F(x) - \Diamond_{\check{y}}^{\hat{y}} F(x)\| \;\le\; \gamma_F \; \max\big[&\|\hat{z} - \hat{y}\| \max(\|x - \check{y}\|, \|x - \check{z}\|), \\ &\|\check{z} - \check{y}\| \max(\|x - \hat{y}\|, \|x - \hat{z}\|)\big] \\ \|\Diamond_{\mathring{z}} F(x) - \Diamond_{\mathring{y}} F(x)\| \quad\le\quad \gamma_F \quad &\|\mathring{z} - \mathring{y}\| \max(\|x - \mathring{y}\|, \|x - \mathring{z}\|) \end{aligned}$$

3 Generalized trapezoidal rule

For the generalization of the trapezoidal rule the linearization mode of choice is the secant mode. The construction largely follows the classical case (see, e.g. [Atk89]). We know that by the fundamental theorem of calculus it holds

$$x(h) - x(0) = \int_0^h F(x(t))\mathrm{d}t \ .$$

Define $\hat{x} := x(h)$ and $\check{x} := x(0)$. We then have:

$$\hat{x} - \check{x} = \int_0^h F(x(t))\mathrm{d}t = h \int_{-\frac{1}{2}}^{\frac{1}{2}} F\left(x\left(\tfrac{h}{2} + \tau h\right)\right) \mathrm{d}\tau \ .$$

Approximating $x(t)$ by the secant

$$h \int_{-\frac{1}{2}}^{\frac{1}{2}} F\left(x\left(\tfrac{h}{2} + \tau h\right)\right) \mathrm{d}\tau = h \int_{-\frac{1}{2}}^{\frac{1}{2}} F\left(\left(\tfrac{1}{2} - \tau\right)\check{x} + \left(\tfrac{1}{2} + \tau\right)\hat{x}\right) \mathrm{d}\tau + \mathcal{O}(h^3)$$

and the latter expression – in contrast to the construction of the classical trapezoidal rule – by its piecewise linearization, we get

$$h \int_{-\frac{1}{2}}^{\frac{1}{2}} F\left(\left(\tfrac{1}{2} - \tau\right)\check{x} + \left(\tfrac{1}{2} + \tau\right)\hat{x}\right) \mathrm{d}\tau = h \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_{\check{x}}^{\hat{x}} F(\mathring{x} + (\hat{x} - \check{x})t)\mathrm{d}\tau + \mathcal{O}(h^3) \ ,$$

where equality holds as a consequence of Proposition 2.1 ii). This yields the following defining equation, which was introduced by Griewank in [Gri13]:

$$\hat{x} - \check{x} = h \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_{\check{x}}^{\hat{x}} F(\mathring{x} + (\hat{x} - \check{x})t)\mathrm{d}t \ , \tag{6}$$

where $\check{x}$ is the current and $\hat{x}$ the next point on the numerical trajectory, c.f. [Gri13, p. 21].

This construction offers two major benefits: Firstly, it has the desired property of having a third order local truncation error, even when integrating through a kink. Secondly, it is consistent with the classical trapezoidal rule in the sense that in case of a smooth function F the generalized formula reduces to the classical one and thus represents a proper extension of the

classical theory [Gri13, S. 21]. It then holds:

$$\begin{aligned}\hat{x} - \check{x} &= h \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_{\check{x}}^{\hat{x}} F(\mathring{x} + (\hat{x} - \check{x})t)\mathrm{d}t = h \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\mathring{F} + \Delta F(\hat{x}, \check{x}; (\hat{x} - \check{x})t)\right] \mathrm{d}t \\ &= h \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\mathring{F} + [F(\hat{x}) - F(\check{x})]\, t\right] \mathrm{d}t = h\mathring{F} \ ,\end{aligned}$$

where $\mathring{F} = \frac{1}{2}\left[F(\hat{x}) + F(\check{x})\right]$. To simplify the equality (6) (which also yields a simplification of next sections' convergence proof), we assume without loss of generality that our current point is the initial value $\check{x} := x(0) = 0$. We then get $\check{x} = 0, x := \hat{x}$ and $\mathring{x} = x/2$ so that the above formula simplifies as follows [Gri13, S. 22]:

$$x = h \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_0^x F(\tfrac{x}{2} + xt)\mathrm{d}t =: hG(x) \ . \tag{7}$$

A generalized midpoint rule can be derived in an analogous fashion (cf. [Gri13, Section 5.2]):

$$\hat{x} - \check{x} = h \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_{\mathring{x}}^{\mathring{x}} F(\mathring{x} + (\hat{x} - \check{x})t)\mathrm{d}t \ .$$

However, while many of the results derived for the trapezoidal rule hold for the midpoint rule as well, the latter does have certain practical disadvantages. For example, the error estimator developed below cannot be applied to it. We thus limit our attention to the trapezoidal rule.

Convergence results

It was shown in [Gri13, Section 5.2] that the generalized midpoint rule has a global accuracy of order two (with respect to the step size h). We will now prove the analogous result for the generalized trapezoidal rule, which has also been stated in [Gri13, Section 5.2]. Note that the arguments we employ for this are mostly similar to those used in the aforementioned reference.

Theorem 3.1 (Griewank)**.** *Suppose, F is piecewise composite differentiable in the sense defined above and Lipschitz continuous in an open neighborhood $\mathcal{D}$ of the origin $\check{x} = 0$. Then there is a bound $\bar{h} > 0$ for the step size, such that for all $h < \bar{h}$ the function $hG(x)$ maps some closed ball $B_\rho(0) \subset \mathcal{D}$, $\rho > 0$, into itself and is contractive. Moreover, the unique fixed point $x_h \in B_\rho(0)$ satisfies the equality*

$$x_h - x(h) \ = \ \mathcal{O}(h^3) \ ,$$

where $x(t)$ is a solution of the differential equation $\dot{x}(t) = F(x(t))$ with initial value $x(0) = 0$.

Proof. F is, by assumption, piecewise composite differentiable and thus locally Lipschitz continuous. Moreover, by Proposition 2.1 we know that the piecewise linearization is Lipschitz continuous. Consequently, with Proposition 2.1 there exists a ball $\mathcal{B}_\rho(0)$ about the base point $x_0 = 0$, such that for all $x \in \mathcal{B}_\rho(0)$ there exists a $\beta_F > 0$ for which it holds

$$F(x) - F(0) \leq \beta_F \rho \quad \text{as well as} \quad \|\diamondsuit_0^x F(0) - \diamondsuit_0^x F(x)\| \leq \beta_F \rho \, . \tag{8}$$

Define γ_F as in Proposition 2.1 iii). Then we have

$$\|\diamondsuit_0^{x_1} F(x) - \diamondsuit_0^{x_2} F(x)\| \leq \gamma_F \, \|x_1 - x_2\| \, \|x\| \ . \tag{9}$$

Employing Equation (7), we get:

$$\begin{aligned}
\|G(\tilde{x}) - G(x)\| &= \left\| \int_{-\frac{1}{2}}^{\frac{1}{2}} \diamondsuit_0^{\tilde{x}} F(\tfrac{\tilde{x}}{2} + t\tilde{x}) - \diamondsuit_0^{x} F(\tfrac{x}{2} + tx) \mathrm{d}t \right\| \\
&\leq \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\| \diamondsuit_0^{\tilde{x}} F(\tfrac{\tilde{x}}{2} + t\tilde{x}) - \diamondsuit_0^{x} F(\tfrac{x}{2} + tx) \right\| \mathrm{d}t \\
&\leq \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\| \diamondsuit_0^{\tilde{x}} F(\tfrac{\tilde{x}}{2} + t\tilde{x}) - \diamondsuit_0^{x} F(\tfrac{\tilde{x}}{2} + t\tilde{x}) \right\| + \left\| \diamondsuit_0^{x} F(\tfrac{\tilde{x}}{2} + t\tilde{x}) - \diamondsuit_0^{x} F(\tfrac{x}{2} + tx) \right\| \mathrm{d}t \equiv \psi \, ,
\end{aligned}$$

which, by (9) and (8), gives

$$\begin{aligned}
\psi &\leq \int_{-\frac{1}{2}}^{\frac{1}{2}} \gamma_F \, \|\tilde{x} - x\| \left\| \tfrac{\tilde{x}}{2} + t\tilde{x} \right\| + \beta_F \left\| \frac{\tilde{x}}{2} + t\tilde{x} - \left(\tfrac{x}{2} + tx \right) \right\| \mathrm{d}t \\
&= \|\tilde{x} - x\| \, (\gamma_F \, \|\tilde{x}\| + \beta_F) \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| t + \tfrac{1}{2} \right| \mathrm{d}t = \tfrac{1}{2} (\gamma_F \, \|\tilde{x}\| + \beta_F) \, \|\tilde{x} - x\| \\
&\leq \tfrac{1}{2} (\gamma_F \rho + \beta_F) \, \|\tilde{x} - x\| =: \tilde{\beta} \, \|\tilde{x} - x\|
\end{aligned}$$

where $\tilde{\beta} = \frac{1}{2}(\gamma_F \rho + \beta_F)$ is a Lipschitz constant for $G(x)$. Consequently, $h\tilde{\beta}$ is a Lipschitz constant for $hG(x)$. Therefore $hG(x)$ is a contraction if we can ensure that $h\tilde{\beta} < 1$, as is the case for h sufficiently small. Since we know

$$G(0) = \frac{1}{2} \left[F(0) + F(0) \right] + \int_{-\frac{1}{2}}^{\frac{1}{2}} \Delta F(0, 0; 0) \mathrm{d}t = F(0)$$

and

$$\|hG(x) - hG(0)\| \geq \|hG(x)\| - \|hG(0)\|,$$

it follows that

$$\|hG(x)\| \leq \|hG(0)\| + h\tilde{\beta}\,\|x\| = h\,\|F(0)\| + h\tilde{\beta}\,\|x\| < \rho$$

for h sufficiently small. Hence, $hG(x)$ maps the ball $\mathcal{B}_\rho(0)$ into itself. With this knowledge we can apply Banach's fixed point theorem and get that the fixed point iteration $hG(x)$ has a unique fixed point $x_h \in \mathcal{B}_\rho(0)$.

Now consider the trajectory $x(t)$ of the exact solution of the differential equation $\dot{x}(t) = F(x(t))$, which is in $C^{1,1}$ since F is Lipschitz continuous. We approximate $x(t)$ with the secant $(t+0.5)x(h)$ for $-0.5 \leq t \leq 0.5$. This corresponds to a polynomial interpolation with a first order polynomial. We will estimate the interpolation error using two auxiliary functions, $g(t)$ and $\delta(t)$. Define

$$\begin{aligned} g(t) &:= x((t+0.5)h) - (t+0.5)x(h) - \frac{\frac{1}{4} - t^2}{\frac{1}{4} - \tau^2}\left[x((\tau+0.5)h) - (\tau+0.5)x(h)\right] \\ &=: \delta(t) - \frac{\frac{1}{4} - t^2}{\frac{1}{4} - \tau^2}\delta(\tau)\,, \end{aligned}$$

where $\delta(t) := x((t+0.5)h) - (t+0.5)x(h)$ and for a $\tau \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. By construction this function is in $C^{1,1}$ and has the three roots $-\frac{1}{2}, \frac{1}{2}, \tau$. Hence, its derivative

$$g'(t) = \delta'(t) - \frac{-2t}{\frac{1}{4} - \tau^2}\delta(\tau)$$

has two roots t_1, t_2. For these it holds:

$$\delta'(t_1) = \frac{-2t_1}{\frac{1}{4} - \tau^2}\delta(\tau) \quad \text{and} \quad \delta'(t_2) = \frac{-2t_2}{\frac{1}{4} - \tau^2}\delta(\tau)\,.$$

Then we have

$$\delta'(t_2) - \delta'(t_1) = \frac{2\delta(\tau)}{\frac{1}{4} - \tau^2}(t_1 - t_2)\,.$$

We know about $\delta'(t)$ that $\delta'(t) = x'((t+0.5)h)h - x(h)$ and consequently

$$\begin{aligned} \left\|\delta'(t_2) - \delta'(t_1)\right\| &= \|F(x(t_2+0.5)h)h - F(x(t_1+0.5)h)h\| \\ &= h\,\|F(x(t_1+0.5)h) - F(x(t_2+0.5)h)\| \\ &\leq \beta_F h^2\,\|x(t_1+0.5) - x(t_2+0.5)\| \leq \alpha_F\beta_F h^2\,\|t_1 - t_2\| \end{aligned}$$

since F is Lipschitz continuous. α_F is defined as $\alpha_F = \sup_{x\in\mathcal{K}}|F(x)|$ for a suitable compact set $\mathcal{K} \subset \mathbb{R}^n$. This yields

$$\left\| \frac{\delta'(t_2) - \delta'(t_1)}{t_1 - t_2} \right\| = \left\| \frac{2\delta(\tau)}{\frac{1}{4} - \tau^2} \right\| \leq \alpha_F \beta_F h^2$$

and accordingly

$$\|\delta(\tau)\| \leq \frac{1}{2}\alpha_F \beta_F h^2 \left(\frac{1}{4} - \tau^2 \right) .$$

Hence, the following inequality holds:

$$\|x((t+0.5)h) - (t+0.5)x(h)\| \leq \frac{\alpha_F \beta_F}{2} \left(\frac{1}{4} - t^2 \right) h^2 , \tag{10}$$

where $t \in (-\frac{1}{2}, \frac{1}{2})$. Moreover, as $\dot{x}(t) = F(x(t))$ and by the fundamental theorem of calculus, we know that

$$x(h) = h \int_{-\frac{1}{2}}^{\frac{1}{2}} F(x(t+0.5)h) \mathrm{d}t . \tag{11}$$

Since it holds

$$\begin{aligned} &\left\| h \int_{-\frac{1}{2}}^{\frac{1}{2}} F(x((t+0.5)h)) \mathrm{d}t - h \int_{-\frac{1}{2}}^{\frac{1}{2}} F((t+0.5)x(h)) \mathrm{d}t \right\| \\ &\leq h \int_{-\frac{1}{2}}^{\frac{1}{2}} \|F(x((t+0.5)h)) - F((t+0.5)x(h))\| \, \mathrm{d}t \;\equiv\; \xi \end{aligned}$$

it follows from the Lipschitz continuity of F that

$$\begin{aligned} \xi &\leq h\beta_F \int_{-\frac{1}{2}}^{\frac{1}{2}} \|x((t+0.5)h) - (t+0.5)x(h)\| \, \mathrm{d}t \\ &\overset{(10)}{\leq} h\beta_F \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\alpha_F \beta_F}{2} \left(\frac{1}{4} - t^2 \right) h^2 \mathrm{d}t \\ &= h^3 \frac{\alpha_F \beta_F^2}{12} \in \mathcal{O}(h^3) . \end{aligned}$$

The latter yields

$$\begin{aligned} x(h) &\overset{(11)}{=} h \int_{-\frac{1}{2}}^{\frac{1}{2}} F(x((t+0.5)h)) \mathrm{d}t \\ &= h \int_{-\frac{1}{2}}^{\frac{1}{2}} F((t+0.5)x(h)) \mathrm{d}t + \mathcal{O}(h^3) . \end{aligned}$$

But, reapplying Proposition 2.1 ii), we also get

$$
\begin{aligned}
&\left| \|x(h) - hG(x(h))\| - \left\| x(h) - h\int_{-\frac{1}{2}}^{\frac{1}{2}} F((t+0.5)x(h))\mathrm{d}t \right\| \right| \\
&\leq h \left\| G(x(h)) - \int_{-\frac{1}{2}}^{\frac{1}{2}} F((t+0.5)x(h))\mathrm{d}t \right\| \\
&\overset{(7)}{=} h \left\| \int_{-\frac{1}{2}}^{\frac{1}{2}} \Diamond_0^{x(h)} F((t+0.5)x(h))\mathrm{d}t - \int_{-\frac{1}{2}}^{\frac{1}{2}} F((t+0.5)x(h))\mathrm{d}t \right\| \\
&\leq h \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\| \Diamond_0^{x(h)} F((t+0.5)x(h)) - F((t+0.5)x(h)) \right\| \mathrm{d}t \\
&\overset{\text{P. 2.1}ii)}{\leq} h\tfrac{1}{2}\gamma_F \int_{-\frac{1}{2}}^{\frac{1}{2}} \|(t+0.5)x(h) - 0\| \, \|(t+0.5)x(h) - x(h)\| \, \mathrm{d}t \\
&= h\tfrac{1}{2}\gamma_F \int_{-\frac{1}{2}}^{\frac{1}{2}} \|(t+0.5)x(h)\| \, \|(t-0.5)x(h)\| \, \mathrm{d}t \\
&= h\tfrac{1}{2}\gamma_F \int_{-\frac{1}{2}}^{\frac{1}{2}} |t+0.5| \, |t-0.5| \, \|x(h)\|^2 \, \mathrm{d}t = h\gamma_F \frac{\|x(h)\|^2}{12} \\
&\leq h\gamma_F \frac{h^2}{12} \in \mathcal{O}(h^3) \, .
\end{aligned}
$$

The last inequality holds, since $x(h) \in \mathcal{B}_\rho(0)$. This implies

$$\|x(h) - hG(x(h))\| \in \mathcal{O}(h^3) \, .$$

Since $\tilde{\beta}h$ is a Lipschitz constant of $hG(x)$, the Banach fixed point theorem provides the following a priori estimate for the distance to the fixed point x_h

$$\|x(h) - x_h\| \leq \frac{\|hG(x(h)) - x(h)\|}{1 - \tilde{\beta}h} \in \mathcal{O}(h^3) \, .$$

□

4 Properties of the presented methods

We say that the solution of an ODE has finite transition if the solution trajectory intersects the nondifferentiabilities of the right hand side in at most a finite point set. This has an impact on the overall performance of the considered methods. If said criterion is violated the classical trapezoidal rule

drops to first order global accuracy. This is not the case for the generalized rule, because the third order local truncation error is maintained even on nondifferentiabilities. This leads to the improved order of accuracy of the generalized method on these problems as illustrated in Table 1.

Without Finite Trans.	on smooth parts	on kinks	globally
Classical Rule	$\mathcal{O}(h^3)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h)$
Generalized Rule	$\mathcal{O}(h^3)$	$\mathcal{O}(h^3)$	$\mathcal{O}(h^2)$

Table 1: Comparison of Accuracy (local, local, global)

Of course, most continuous examples do have finite transition. However, even in this case we can still expect a gain as described in the following subsection. Note that finite transition does not require the trajectory to be transversal to the sets of kink locations, i.e. the solution trajectory may be tangential to the set of nondifferentiabilities of F at the point of intersection. For a definition of transversality see, e.g. [dBBCK08, p. 64]. The latter, stronger property is required for efficient event handling either by educated guessing or by computing the roots of switching functions, mainly applied for discontinuous ODEs [EJNT88, ST00]. Switching functions are necessarily singular at tangential transition points.

Romberg Extrapolation

In the following we assume that all solutions have finite transition. It is well known that the local truncation error of the trapezoidal rule can be improved to order five by applying Richardson extrapolation once for a sufficiently smooth function F. We will refer to this form of a Romberg's method as Romberg extrapolation. Hence its maximal order that can be expected for the global error is four. However, if the aforementioned F is only piecewise differentiable, then the order collapses on the kinks, posing an upper bound for the global error. Thus the overall global accuracy in the case of Romberg extrapolation is determined by the respective behavior of the investigated method on the nondifferentiablities of F. Here the generalized method should be superior, since its accuracy only collapses to an error of order three, as opposed to order two for the classical method. This is the error that the respective methods would achieve without extrapolation. It is lower for the classical method, because the linear approximation used in its construction is only of first order on the kinks as opposed to second order for the piecewise linear approximation of the generalized method. This is summarized in Table 2.

With Finite Transition	on smooth parts	on kinks	globally
Classical Rule	$\mathcal{O}(h^3)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$
Generalized Rule	$\mathcal{O}(h^3)$	$\mathcal{O}(h^3)$	$\mathcal{O}(h^2)$
Class. w/ Romberg	$\mathcal{O}(h^5)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$
Gen. w/ Romberg	$\mathcal{O}(h^5)$	$\mathcal{O}(h^3)$	$\mathcal{O}(h^3)$

Table 2: Accuracy including Romberg Extrapolation (local, local, global)

One might have hoped that extrapolation would also yield a local error of $\mathcal{O}(h^5)$ on kinks. The following, easily verifiable example shows that this is not the case.

Lemma 4.1. *The analytical solution to the nonsmooth ODE*

$$\dot{x} = a|x| + bx + 1 = \begin{cases} (b+a)x + 1 & x \geq 0 \\ (b-a)x + 1 & x < 0 \end{cases}$$

with $x(0) = 0$ *is*

$$x_-(t) = \frac{e^{t(b-a)} - 1}{b-a} \quad \textit{and} \quad x_+(t) = \frac{e^{t(b+a)} - 1}{b+a} \, .$$

If now we set $a = \frac{9}{4}$ and $b = -\frac{5}{4}$, we can calculate the local truncation error for a single step from $\breve{x} = -0.5$ with step size $h = 1$ over the kink at $x = 0$. For the classical method it amounts to $\frac{27h^2}{64} + \frac{677h^3}{1536} + \mathcal{O}(h^4)$ and for the generalized method to $\frac{139h^3}{3072} + \mathcal{O}(h^4)$. Using Romberg extrapolation does not improve the order of the error. In this case we get $\frac{3h^2}{64} + \frac{51h^3}{512} + \mathcal{O}(h^4)$ for the classical method and $\frac{9h^3}{1024} + \mathcal{O}(h^4)$ for the generalized method. Consequently, as opposed to the smooth case, Romberg extrapolation only yields a third order global convergence instead of the usual $\mathcal{O}(h^4)$.

Remark on Geometric Integration

Among ODE integration methods, those which allow for the preservation of certain geometric properties, especially energy preservation and symplecticness, play an important role in current research. The presented method is part of the latter category for piecewise linear Hamiltonian systems. To show this, we note that the piecewise linearization of a piecewise linear function is the function itself. Hence, for a piecewise linear right hand side, the formula for the generalized trapezoidal rule simplifies as follows:

$$\hat{x} - \check{x} = h \int_{-\frac{1}{2}}^{\frac{1}{2}} F\left(\frac{\hat{x} + \check{x}}{2} + t(\hat{x} - \check{x})\right) \mathrm{d}t \ .$$

In [Qui08] the concept of a so-called average vector field method (AVF) is introduced in terms of the following formula:

$$\hat{x} - \check{x} = h \int_0^1 F((1-s)\check{x} + s\hat{x}) \mathrm{d}s \ .$$

An AVF is energy preserving on Hamiltonian systems. However, since for this method an exact integration of the right hand side is necessary, there only exists a straightforward implementation for linear systems. But the generalized trapezoidal rule performs an exact integration of the right hand side in case F is piecewise linear. Thus it is energy preserving on piecewise linear Hamiltonians.

5 Dense Output by Polynomial Interpolation

It is well known that given a Runge-Kutta method for a smooth system one can derive a dense output as approximation of the solution trajectory, e.g. via Hermite interpolation [HNW93, p. 188ff]. On a single integration step, this takes the form of a quadratic polynomial $p\colon [0,h] \to \mathbb{R}^n$ with a third order approximation error. For the trapezoidal rule it is given by

$$p(t) = \check{x} + \int_0^t F(\check{x}) + \tfrac{\tau}{h}(F(\hat{x}) - F(\check{x})) \ \mathrm{d}\tau \quad \text{for } t \in [0,h] \ .$$

This polynomial is tangential to the numerical trajectory in the sense that its values at $t = 0$ and $t = h$ equal $\check{x}$ and $\hat{x}$, respectively, and its slope matches the vector field of the numerical solution in these points. The integral can be evaluated and gives an explicit formula

$$p(t) = \frac{F(\hat{x}) - F(\check{x})}{2h} t^2 + F(\check{x})t + \check{x} \ .$$

However, in the case of a step through a kink, the linear approximation of F used in the trapezoidal rule is only a first order approximation. Thus the above polynomial is only a second order approximation of the trajectory, which is not sufficient for certain applications like the construction of an error estimator that we pursue in Section 6. Fortunately, the generalized trapezoidal rule allows for the construction of such a polynomial with the

desired properties for nonsmooth functions. It is derived in an analogous way and takes the form:

$$p(t) \;=\; \check{x} + \int_0^t \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + \tfrac{\tau}{h}(\hat{x} - \check{x}))\mathrm{d}\tau \quad \text{for } t \in [0, h] \; .$$

Of course it is now a piecewise quadratic function which consists of some p_i with

$$p_i(t) = p|_{[h\tau_i, h\tau_{i+1}]}\,(h\tau_i + t) \;=\; a_i t^2 + b_i t + c_i \; .$$

Here $h\tau_i$ is the length of the step to the next kink, or to $\hat{x}$ if there is no kink. These values are contained in the piecewise linearization of F and are calculated during the integration. We also need the intermediate values $\check{x}_{\rightarrow,i}$ of the numerical trajectory at the kinks for which it holds:

$$\check{x}_{\rightarrow,i} = \check{x} + h \int_0^{\tau_i} \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + t(\hat{x} - \check{x}))\mathrm{d}t \; .$$

They are already calculated as well, since the integral from $\check{x}$ to $\hat{x}$ is simply the sum of the intermediate values. If there are k kinks in the observed time step, we have $\tau_0 = 0$ and $\tau_{k+1} = 1$, such that $\check{x}_{\rightarrow,k+1} = \hat{x}$. Consequently, the derivatives $\dot{\check{x}}_{\rightarrow,i}$ are given by $\dot{\check{x}}_{\rightarrow,i} = F(\check{x}_{\rightarrow,i})$. We can now derive the coefficients of the interpolants p_i, since we know that:

$$\begin{aligned}
p_i(0) = \check{x}_{\rightarrow,i} \quad &\Longrightarrow \quad c_i = \check{x}_{\rightarrow,i} \; , \\
p_i'(0) = \dot{\check{x}}_{\rightarrow,i} \quad &\Longrightarrow \quad b_i = \dot{\check{x}}_{\rightarrow,i} \; , \\
p_i'(h(\tau_{i+1} - \tau_i)) = \dot{\check{x}}_{\rightarrow,i+1} \Longrightarrow \quad & a_i = \frac{\dot{\check{x}}_{\rightarrow,i+1} - \dot{\check{x}}_{\rightarrow,i}}{2h(\tau_{i+1} - \tau_i)} \; .
\end{aligned}$$

This means p is given by

$$p\left(\tilde{t}\right) = p(h\tau_i + t) = p_i(t) = \frac{\dot{\check{x}}_{\rightarrow,i+1} - \dot{\check{x}}_{\rightarrow,i}}{2h(\tau_{i+1} - \tau_i)} t^2 + \dot{\check{x}}_{\rightarrow,i} t + \check{x}_{\rightarrow,i} \quad \text{for } \tilde{t} \in [h\tau_i, h\tau_{i+1}], i \in \{0, \ldots, k\} \; .$$

This polynomial has correct values and slopes at $\check{x}, \hat{x}$ and all the kinks. Accordingly, it is a third order approximation on the intervals between consecutive kinks or $\check{x}$ or $\hat{x}$, respectively and thus everywhere. It is depicted in Figure 2 for the first example in Section 7, defined by Equation (12).

6 Error Estimation and Time-Stepping

In this section we will construct an error estimator for the local truncation error of the generalized trapezoidal rule.

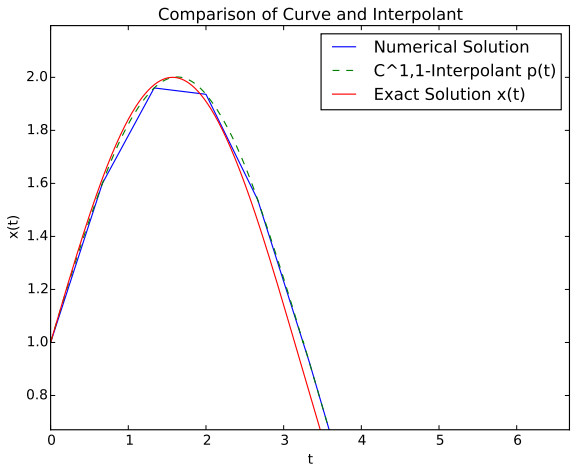


Figure 2: Dense Output, first num. Example, $h = \frac{2}{3}$

Proposition 6.1 (Error estimator)**.** *The error estimator has the form*

$$\begin{aligned}&\|x(h) - x_h\| \leq \tfrac{1}{12} h \gamma_F \left\|\hat{x} - \check{x}\right\|^2 \\ &+ \beta_F \sum_{i=0}^{k} \int_0^{h(\tau_{i+1}-\tau_i)} \left\| \frac{\dot{\check{x}}_{\rightarrow,i+1} - \dot{\check{x}}_{\rightarrow,i}}{2h(\tau_{i+1}-\tau_i)} t^2 + \left(\dot{\check{x}}_{\rightarrow,i} - \frac{\hat{x} - \check{x}}{h} \right) t + \check{x}_{\rightarrow,i} - \check{x} - \tau_i(\hat{x} - \check{x}) \right\| \mathrm{d}t \,.\end{aligned}$$

Note that for the evaluation of this formula the integral over the absolute value of a quadratic function has to be calculated. In general, this necessitates the computation of the roots of the polynomial.

We start the derivation of the error estimator by splitting up the truncation error into two components:

$$\begin{aligned}\|x(h) - x_h\| &= \left\| \int_0^h F(x(t))\mathrm{d}t - \int_0^h \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x}))\mathrm{d}t \right\| \\ &\leq \int_0^h \left\| F(x(t)) - \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) \right\| \mathrm{d}t \\ &\leq \int_0^h \left\| F(x(t)) - F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) \right\| \\ &\quad + \left\| F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) - \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) \right\| \mathrm{d}t \,.\end{aligned}$$

The left term of the last expression can be bounded, using the Lipschitz constant β_F from Proposition 2.1:

$$\int_0^h \left\| F(x(t)) - F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) \right\| \mathrm{d}t \leq \beta_F \int_0^h \left\| x(t) - \check{x} - \tfrac{t}{h}(\hat{x} - \check{x}) \right\| \mathrm{d}t \; .$$

Since the analytical solution trajectory is unknown, we approximate it with the piecewise quadratic interpolation polynomial constructed above, whose approximation is of order $\mathcal{O}(h^3)$, which does not decrease the overall approximation error of order two.

$$\begin{aligned} \beta_F \int_0^h \left\| x(t) - \check{x} - \tfrac{t}{h}(\hat{x} - \check{x}) \right\| \mathrm{d}t &= \beta_F \int_0^h \left\| p(t) - \check{x} - \tfrac{t}{h}(\hat{x} - \check{x}) + \mathcal{O}(h^3) \right\| \mathrm{d}t \\ &= \beta_F \int_0^h \left\| p(t) - \check{x} - \tfrac{t}{h}(\hat{x} - \check{x}) \right\| \mathrm{d}t + \mathcal{O}(h^4) \; . \end{aligned}$$

This integral can be split up into the sum of the integrals from kink to kink:

$$\beta_F \int_0^h \left\| p(t) - \check{x} - \tfrac{t}{h}(\hat{x} - \check{x}) \right\| \mathrm{d}t = \beta_F \sum_{i=0}^{k} \int_0^{h(\tau_{i+1} - \tau_i)} \left\| p_i(t) - \check{x} - \tfrac{h\tau_i + t}{h}(\hat{x} - \check{x}) \right\| \mathrm{d}t$$

$$= \beta_F \sum_{i=0}^{k} \int_0^{h(\tau_{i+1} - \tau_i)} \left\| \frac{\dot{\check{x}}_{\rightarrow,i+1} - \dot{\check{x}}_{\rightarrow,i}}{2h(\tau_{i+1} - \tau_i)} t^2 + \left(\dot{\check{x}}_{\rightarrow,i} - \frac{\hat{x} - \check{x}}{h} \right) t + \check{x}_{\rightarrow,i} - \check{x} - \tau_i(\hat{x} - \check{x}) \right\| \mathrm{d}t \; ,$$

which is the first part of the error estimator. The second part can be bounded, using Proposition 2.1 ii):

$$\begin{aligned} &\int_0^h \left\| F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) - \Diamond_{\check{x}}^{\hat{x}} F(\check{x} + \tfrac{t}{h}(\hat{x} - \check{x})) \right\| \mathrm{d}t \\ &\leq \tfrac{1}{2}\gamma_F \int_0^h \left\| \tfrac{t}{h}(\hat{x} - \check{x}) \right\| \left\| (\tfrac{t}{h} - 1)(\hat{x} - \check{x}) \right\| \mathrm{d}t \\ &= \tfrac{1}{2}\gamma_F \left\| \hat{x} - \check{x} \right\|^2 \int_0^h \left| \tfrac{t^2}{h^2} - \tfrac{t}{h} \right| \mathrm{d}t = \tfrac{1}{12} h \gamma_F \left\| \hat{x} - \check{x} \right\|^2 \; . \end{aligned}$$

Hence, the overall error estimator has the form stated in Proposition 6.1. With this formula in hand, a step size control can be implemented, just as in the classical case.

7 Numerical Examples

Rolling Stone

This example tracks a point moving without friction on a convex surface representing an idealized rolling stone. It can be considered as a harmonic

oscillator, provided the surface is parabolic. We modify this parabola by inserting a planar section in the interval $[-1, 1]$. This yields the curve

$$V(x) = \begin{cases} \frac{1}{2}(1+x)^2, & x \leq -1 \\ 0\,, & -1 < x < 1\,. \\ \frac{1}{2}(1-x)^2, & 1 \leq x \end{cases}$$

The derivative of V defining the acceleration $\ddot{x}$ of the mass is piecewise linear and given by

$$-V'(x) = \min(\max(-1-x, 0), 1-x) = -x - |x-1|/2 + |x+1|/2$$

which yields the ODE $\ddot{x} = -V'(x)$. The analytic solution for $x(0) = 1$, $\dot{x}(0) = 1$ is $(2\pi+4)$-periodic and given by

$$x(t) = \begin{cases} 1 + \sin(t) & 0 \leq t \leq \pi \\ 1 - (t-\pi) & \pi \leq t < \pi + 2 \\ -1 - \sin(2-t) & \pi + 2 \leq t < 2\pi + 2 \\ t - 3 - 2\pi & 2\pi + 2 \leq t < 2\pi + 4 \end{cases}$$

In Figure 3 we depicted V and V' as well as the analytic solution of the ODE. The linear parts are drawn in gray.
From the second order ODE we obtain the first order system

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 - |x_1 - 1|/2 + |x_1 + 1|/2 \end{pmatrix} = F(x)\,. \tag{12}$$

We will consider the initial conditions $x_1(0) = 1$ and $x_2(0) = 1$. As predicted, in Figure 4 we observe a global convergence order of two for both the classical and generalized method. However, in Figure 4 it is clearly visible that, in contrast to the generalized method, the error of the classical method does not decrease monotonically with h. This is a consequence of the generalized method's greater accuracy on the kinks, which is also the reason that extrapolating yields an increased convergence order only for the generalized method.

Due to the simplicity of the example adaptive time-stepping does not improve the solution significantly. We thus omit presenting the associated results. But as a frictionless mechanical system it is energy preserving. As a piecewise linear Hamiltonian system it fulfills the requirements for the generalized method to correctly preserve this energy. Accordingly, the total energy

$$V(x(t)) + \tfrac{1}{2}\dot{x}(t)^2$$

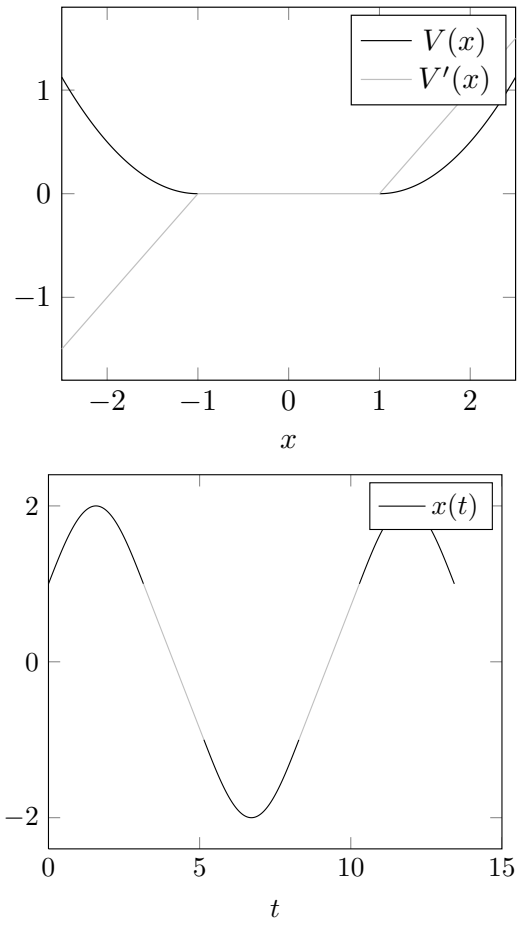


Figure 3: Visualization of V and the analytic solution of the ODE

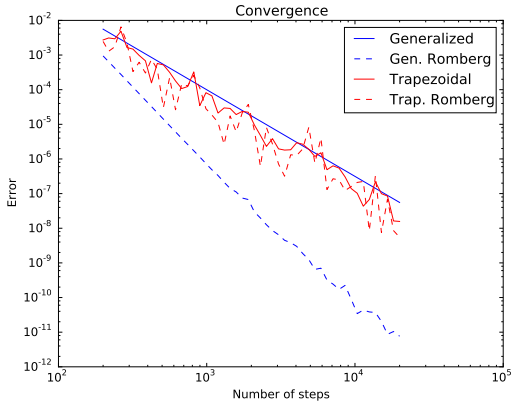


Figure 4: Convergence global Error Rolling Stone

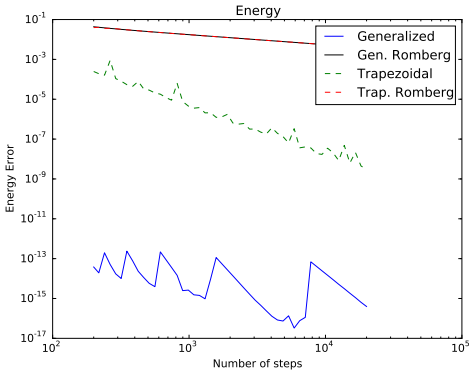
must stay constant at its initial value $V(x_0) + \frac{1}{2}y_0^2 = \frac{1}{2}$. Hence, we consider the total energy variation over all N time steps using the formula

$$\text{Energy error} = \left[\sum_{i=1}^{N} \left(V(x_{1,i}) + \tfrac{1}{2}(x_{2,i})^2 - \tfrac{1}{2}\right)^2\right]^{\frac{1}{2}}.$$

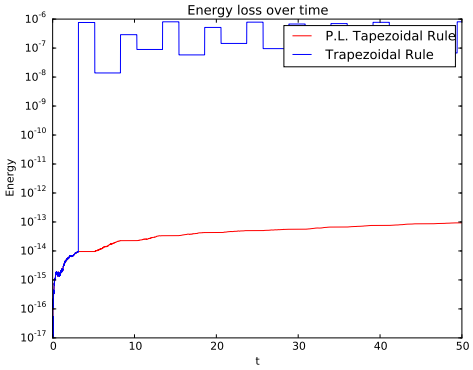
We observe in Figure 5 (a) that the energy error of the generalized method is at the level of machine precision as expected, whereas the classical method is not energy preserving on the kinks, as can be seen in Figure 5 (b), and consequently does not preserve the systems' energy globally, either. However, unfortunately we can also see that Romberg extrapolation does not leave the energy preservation property of the generalized method intact. That is, both methods perform equally poor in the situation at hand. A possible restoration of this property remains to be investigated.

Diode LC-Circuit

The second example closer resembles problems arising in actual applications. We take a simple LC-circuit and replace the resistor with a diode, thus providing an element which causes a nondifferentiable impact in the equations describing the system. Figure 6 depicts the circuit. It is modeled by the following system of ODE's, where x_1 represents time, x_2 represents the charge



(a) Convergence of Energy Error



(b) Energy Loss over Time, $h = 10^{-4}$

Figure 5: Energy Preservation

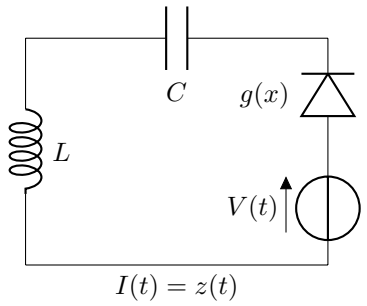


Figure 6: Circuit Diagram

(at the capacitor) and x_3 represents the electric current in the circuit.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = F(\mathbf{x}) = \begin{pmatrix} 1 \\ x_3 \\ -\big(x_2 - CV(x_1) + g(Cx_3)\big)\frac{1}{LC} \end{pmatrix} \tag{13}$$

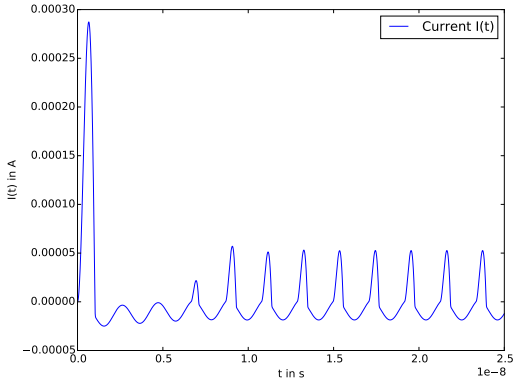
Here L and C are the inductance and the capacitance, respectively, of the corresponding elements L and C in Figure 6. Furthermore, $V(x_1) = \sin(\omega\ x_1)$ is the forcing current and $g(z)$ models the diode (for small currents) in the piecewise linear form

$$g(z) = \frac{z + |z|}{2\alpha} + \frac{z - |z|}{2\beta} = \begin{cases} \frac{z}{\alpha} & \text{if } z \geq 0 \\ \frac{z}{\beta} & \text{if } z < 0. \end{cases}$$

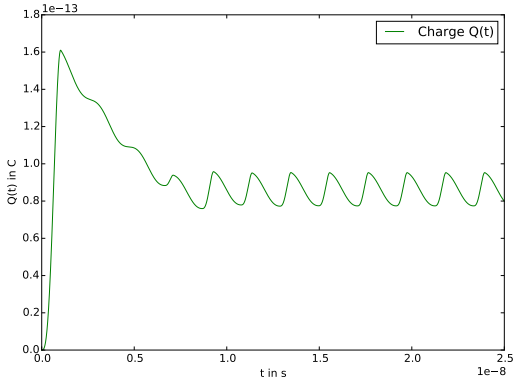
We choose a set of constants that resembles those occurring in actual electric circuits:

$$L = 10^{-6}\ \text{H},\ C = 10^{-13}\ \text{F},\ \omega = 3 \cdot 10^9\ \text{Hz},\ \alpha = 2\ \Omega^{-1},\ \beta = 0.00001\ \Omega^{-1}\ .$$

Moreover, we consider the initial conditions $x_1(0) = x_2(0) = x_3(0) = 0$. The result of the numerical integration of (13) employing any of the considered methods for $h = 10^{-4}$ is depicted in Figure 7, (a) and (b). As one can see, the capacitor is initially charged over one cycle and discharged over a few more, before the solution adopts a periodic behavior. The solution trajectory changes its behavior every time the current changes its sign. As presented in Figure 8, both methods display second order convergence, while in case of Romberg extrapolation the generalized method gains an order and the classical does not, just like in the previous example. Also, we observe that adaptive time-stepping results in similar improvements to the convergence constant for the generalized method as for the classical method.



(a) Current



(b) Charge

Figure 7: Solution of the ODE System

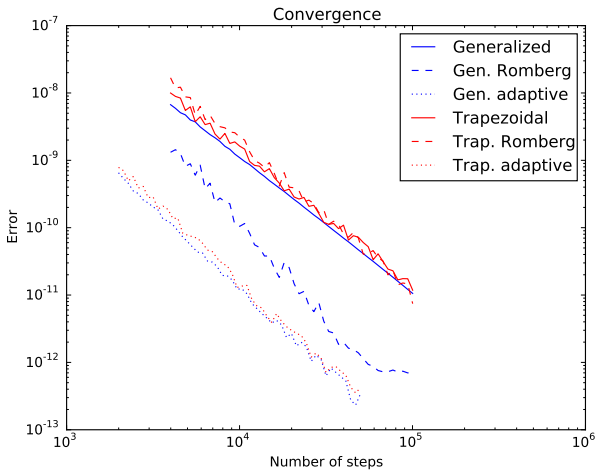


Figure 8: Convergence Global Error Diode Example

8 Final Remarks

As mentioned before, the energy preservation is lost after extrapolation, as observable in Figure 5 a). This can be partly explained by the loss of reversibility for the numerical integrator with extrapolation. It has to be investigated further, if this property can be restored. Furthermore, the current Lipschitz constants from Proposition 2.1 are in some cases very conservative overestimations. It is desirable to sharpen these bounds. Also, one should investigate the possibility of automated scaling of the error norm using information from the structure of the piecewise linearization to reflect the dimensions of the components of the error. A long term goal is the extension of the method to piecewise smooth functions that are not continuous. In this publication, the question of efficient implementation and computational cost of the described method has not been studied in detail. It will be the subject of further investigations. Each inner iteration of the method requires the solution of a piecewise linear system, e.g. by the solvers proposed in [Rad16, SGRB14, GBRS15]. We intend to deliver an efficient implementation for an integrated framework of piecewise linearization and ODE as well as equation solving in subsequent publications.

Acknowledgements

The work for the article has been partially conducted within the Research Campus MODAL funded by the German Federal Ministry of Education and Research (BMBF) (fund number 05M14ZAM).

References

[Atk89] K. Atkinson. *An Introduction to Numerical Analysis (2nd ed.)*. John Wiley & Sons, New York, 1989.

[Cla83] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. Wiley-Interscience, 1983.

[dBBCK08] M. di Bernardo, C.J. Budd, A.R. Champneys, and P. Kowalczyk. *Piecewise-smooth Dynamical Systems*. Applied Mathematical Sciences. Springer, 2008.

[EJNT88] W.H. Enright, K.R. Jackson, S.P. Nørsett, and P.G. Thomsen. Effective solution of discontinuous ivps using a runge-kutta formula pair with interpolants. *Applied Mathematics and Computation*, 27:313–335, 1988.

[GBRS15] A. Griewank, J.U. Bernt, M. Radons, and T. Streubel. Solving piecewise linear systems in abs-normal form. *Linear Algebra and its Applications*, 471(0):500 – 530, 2015.

[Gri13] A. Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013.

[GSHR] A. Griewank, T. Streubel, R. Hasenfelder, and M. Radons. Piecewise linear secant approximation via algorithmic piecewise differentiation. *Submitted 2016*.

[GW08] A. Griewank and A. Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), 2008.

[HNW93] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I.* Springer Series in Computational Mathematics. Springer, 1993.

[Nau12] Uwe Naumann. *The Art of Differentiating Computer Programs: An Introduction to Algorithmic Differentiation.* SIAM, 2012.

[Qui08] D.I. Quispel, G.R.W.; McLaren. A new class of energy-preserving numerical integration methods. *Journal of Physics A: Mathematical and Theoretical*, 41, 2008.

[Rad16] M. Radons. Direct solution of piecewise linear systems. *Theoretical Computer Science*, 626:97–109, 2016.

[Sch12] S. Scholtes. *Introduction to Piecewise Differentiable Equations.* SpringerBriefs in optimization. Springer New York, 2012.

[SGRB14] T. Streubel, A. Griewank, M. Radons, and J.U. Bernt. Representation and analysis of piecewise linear functions in absnormal form. *Proc. of the IFIP TC 7*, pages 323–332, 2014.

[ST00] L.F. Shampine and S. Thomson. Event location for ordinary differential equations. *Computers & Mathematics with Applications*, 39:43–54, 2000.