AN EXPONENTIAL TIME-DIFFERENCING METHOD FOR MONOTONIC RELAXATION SYSTEMS

PEDER AURSAND[A,D], STEINAR EVJE[B], TORE FLÅTTEN[C,D],
KNUT ERIK TEIGEN GILJARHUS[D] AND SVEND TOLLAK MUNKEJORD[D,E]

Abstract. We present first and second-order accurate exponential time differencing methods for a special class of stiff ODEs, denoted as *monotonic relaxation ODEs*. Some desirable accuracy and robustness properties of our methods are established. In particular, we prove a strong form of stability denoted as *monotonic asymptotic stability*, guaranteeing that no overshoots of the equilibrium value are possible. This is motivated by the desire to avoid spurious unphysical values that could crash a large simulation.

We present a simple numerical example, demonstrating the potential for increased accuracy and robustness compared to established Runge–Kutta and exponential methods. Through operator splitting, an application to granular-gas flow is provided.

subject classification. 65L04, 65L06, 65M08, 76T25

key words. exponential integrators, relaxation, stiff systems

1. Introduction

We are interested in numerical methods for stiff relaxation systems in the form

$$\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} = \frac{1}{\epsilon}\boldsymbol{S}(\boldsymbol{V}), \tag{1}$$

to be solved for the unknown vector $\boldsymbol{V}$. Herein, the effect of the *relaxation source term* $\boldsymbol{S}(\boldsymbol{V})$ is to drive the system towards some local equilibrium value $\boldsymbol{V}^{\mathrm{eq}}$. The parameter ϵ represents a characteristic *relaxation time* towards equilibrium.

Our motivation for studying such systems is their relevance for more general hyperbolic relaxation systems in the form

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial \boldsymbol{F}(\boldsymbol{U})}{\partial x} = \frac{1}{\epsilon}\boldsymbol{R}(\boldsymbol{U}), \tag{2}$$

as analysed in detail by Chen et al. [4]. The parameter ϵ is typically small, imposing a high degree of stiffness in the system (2).

Date: February 24, 2014.

[A]Dept. of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway.

[B]Dept. of Petroleum Engineering, University of Stavanger (UiS), NO-4036 Stavanger, Norway.

[C]SINTEF Materials and Chemistry, P. O. Box 4760 Sluppen, NO-7465 Trondheim, Norway.

[D]SINTEF Energy Research, P.O. Box 4761 Sluppen, NO-7465 Trondheim, Norway.

Email: peder.aursand@math.ntnu.no, Steinar.Evje@uis.no, Tore.Flatten@sintef.no, Knut.Erik.Giljarhus@sintef.no, stm@pvv.org.

[E]Corresponding author.

A popular approach towards solving stiff systems in the form (1) has been the use of *exponential integrators* [5, 12, 22]. Such methods are motivated in part by computational efficiency considerations [13]; without sacrificing high-order accuracy, one gets rid of the severe restriction on the time step commonly associated with explicit methods for stiff problems. The main idea behind such methods consists of splitting the source term into a linear and a nonlinear part as follows:

$$\frac{1}{\epsilon}\boldsymbol{S}(\boldsymbol{V}) = \boldsymbol{L}\boldsymbol{V} + \boldsymbol{N}(\boldsymbol{V}), \tag{3}$$

where $\boldsymbol{L}$ is a constant matrix. Ideally, the stiffness of the system (1) should be associated with the linear part, which may be solved exactly through the matrix exponential. Coupled to this, the non-linear part $\boldsymbol{N}(\boldsymbol{V})$ is solved by standard Runge–Kutta methods.

In this paper, we wish to emphasize another aspect of methods based on exponential decay; the potential for strong robustness in the sense that the numerical solution is bounded with no restriction on the time step. In particular, one may use such methods to ensure that the relaxation step does not introduce unphysical solutions such as vacuum or negative-density states.

To achieve this, we here present what seems to us a slightly original twist to the idea of exponential integrators. Instead of viewing the exponential integration step as the *exact* solution to a linear sub-problem as given by the splitting (3), we interpret the exponential integration as a *numerical approximation* to the original nonlinear problem, and this approximation is nevertheless accurate to a certain order in the time step. This change of perspective leads to a slightly different formulation, and allows us to construct consistent methods that *by design* guarantee that the equilibrium solution cannot be exceeded. Although this leads to a high degree of robustness and accuracy in the stiff limit, the error of our proposed method nevertheless formally depends on the stiffness of the system.

If the numerical solution is bounded by the equilibrium value, consistency requires the same bound to hold also for the exact mathematical solution. Therefore, we will limit our investigations in this paper to what we denote as *monotonic* equations in the form (1), as defined more precisely in Section 2. This restricts the class of systems where our methods are applicable, but in particular includes many relaxation processes of practical interest within the context of (2).

This paper is organized as follows. In Section 2, we present the exponential integration technique which is the topic of this paper. First and second-order versions are provided. We also prove the following.

(i) The methods are stable in the strong sense that no numerical overshoots of the equilibrium value are possible.
(ii) The error is of second order in perturbations from the equilibrium if the source term decays linearly to zero.

Technical details needed for these proofs are given in Appendix A.

In Section 3, we briefly review hyperbolic relaxation systems in the form (2), and some known challenges associated with developing numerical methods for such systems. In this context, we discuss the potential applicability of the methods derived in Section 2.

In Section 4, some numerical examples are presented. In Section 4.1, we illustrate the main strength of our methods; they respect the monotonicity of the original equation with no restrictions on the time step. This example also demonstrates how

standard Runge–Kutta methods and a classical exponential integrator may fail to possess this property.

This high degree of numerical stability would be desirable when solving more general hyperbolic relaxation systems in the form (2). Therefore, in Section 4.2, we present some preliminary investigations on applying our methods to such systems. In particular, we consider a numerical benchmark case known from the literature; a model for *granular-gas flow* as investigated by Serna and Marquina [33]. These initial tests seem to compare satisfactorily to results previously reported in the literature, indicating that our proposed methods may be worthy of further investigation.

Finally, in Section 5 we summarize our results and discuss some directions for further work.

2. Monotonically Asymptotic Exponential Integration

For the purposes of this paper, we make the following definition.

Definition 1. *Consider the equation*

$$\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t} = \frac{1}{\epsilon}\boldsymbol{S}(\boldsymbol{V}), \qquad \boldsymbol{V} \in \mathcal{D} \subseteq \mathbb{R}^N, \qquad \boldsymbol{V}(0) = \boldsymbol{V}_0 \in \mathcal{D} \tag{4}$$

where $\boldsymbol{S}(\boldsymbol{V})$ is a C^2 function. The system is said to be a **relaxation ODE** *provided there exists a unique point $\boldsymbol{V}^{\mathrm{eq}} \in \mathcal{D}$ such that*

$$\boldsymbol{S}(\boldsymbol{V}^{\mathrm{eq}}) = 0, \tag{5}$$

and the solution satisfies

$$\lim_{t\to\infty} \boldsymbol{V}(t) = \boldsymbol{V}^{\mathrm{eq}}. \tag{6}$$

2.1. Exponential Integrators.

A classical method for the time integration of systems of stiff differential equations is to use *exponential integrators*. Such methods have for several decades constituted an active field of research [6, 23, 11].

The basic idea behind exponential integrators is a splitting of the source term into a linear and a nonlinear part as follows [5, 13, 16]:

$$\frac{1}{\epsilon}\boldsymbol{S}(\boldsymbol{V}) = \boldsymbol{L}\boldsymbol{V} + \boldsymbol{N}(\boldsymbol{V}), \tag{7}$$

where $\boldsymbol{L}$ is a constant matrix. The linear part can then be solved exactly through application of the matrix exponential. If the stiffness can be associated with the linear term only, i.e. if

$$\boldsymbol{S}(\boldsymbol{V}) = \mathcal{L}\boldsymbol{V} + \epsilon\mathcal{N}(\boldsymbol{V}), \tag{8}$$

where $\mathcal{L}$ and $\mathcal{N}$ are independent of ϵ, such methods have the potential for error bounds that do not depend on ϵ.

Different types of exponential integrators exist. One classical approach is Lawson's method [23, 22], where one starts by performing the variable transformation

$$\boldsymbol{W}(t) = \exp\left(-\frac{t}{\epsilon}\boldsymbol{L}\right)\boldsymbol{V}(t), \tag{9}$$

which may be substituted in (4) to yield

$$\frac{\mathrm{d}\boldsymbol{W}}{\mathrm{d}t} = \frac{1}{\epsilon}\exp\left(-\frac{t}{\epsilon}\boldsymbol{L}\right)\boldsymbol{S}(\boldsymbol{V}(\boldsymbol{W})) - \frac{1}{\epsilon}\boldsymbol{L}\boldsymbol{W}. \tag{10}$$

One then simply solves for $\boldsymbol{W}$ using a standard Runge–Kutta scheme. Other important classes of exponential integrators include exponential Runge–Kutta methods [14], exponential Rosenbrock methods [15] and exponential multistep methods [27]. For a detailed account on different approaches and error analysis, we refer to the recent review by Hochbruck and Ostermann [17].

For stiff problems, exponential integrators allow for larger time steps and improved stability compared to straightforward Runge–Kutta methods. A general theory for constructing high-order versions, applicable to a rather large class of exponential integrators, was presented by Berland et al. [1].

2.2. **Monotonic Relaxation ODEs.** Much of the existing literature on exponential integrators focuses on computational *accuracy* and *efficiency*. Our current method is motivated by the desire to shift the focus more strongly towards numerical *robustness*. Towards this end, we first define a subclass of relaxation ODEs.

Definition 2. *A relaxation ODE in the form* (4) *is said to be a* **monotonic relaxation ODE** *if*

$$V_i'(t)\,(V_i^{\mathrm{eq}} - V_i(t)) > 0 \quad \forall V_i \neq V_i^{\mathrm{eq}} \tag{11}$$

for all $i \in \{1, \ldots, N\}$.

In other words, we denote the system as monotonic if all the components of the solution vector are monotonic functions of time. From (4) and (11) we immediately see that a *necessary* condition for a system in the form (4) to be a monotonic relaxation ODE is that the source term must satisfy

$$S_i(\boldsymbol{V})\,(V_i^{\mathrm{eq}} - V_i) > 0 \quad \forall V_i \neq V_i^{\mathrm{eq}} \tag{12}$$

for all $i \in \{1, \ldots, N\}$.

Proposition 1. *The source term of a monotonic relaxation ODE satisfies*

$$\frac{\partial S_i}{\partial V_j}(\boldsymbol{V}^{\mathrm{eq}}) = 0 \quad \forall i \neq j. \tag{13}$$

Proof. It follows from monotonicity that we must have

$$S_i(\boldsymbol{V}) = 0 \quad \text{if} \quad V_i = V_i^{\mathrm{eq}}, \tag{14}$$

or else (6) and (12) cannot simultaneously hold. □

Within the framework of hyperbolic relaxation systems in the form (2), monotonicity seems to be a rather inclusive restriction. For instance, it is an essential property of scalar relaxation ODEs.

Proposition 2. *All scalar relaxation ODEs are monotonic in the sense of Definition 2.*

Proof. Assume there exists some time $t = \bar{t}$ where $V'(t)$ changes sign. Given that $S(V)$ is a smooth function, we would here have $V(\bar{t}) = V^{\mathrm{eq}}$ which would hold for all $t \geq \bar{t}$. Hence (11) is automatically satisfied. □

If the relaxation processes are fully independent, this property will carry directly over to systems. For instance, the relaxation part of the five-equation two-phase flow model investigated by Munkejord [25], describing simultaneous volume and momentum transfer, consists of independent relaxation processes and is monotonic in the sense of Definition 2.

We will consider a concrete example of a nonlinear, coupled monotonic relaxation system in Section 4.1. In general, however, strongly coupled relaxation systems cannot be expected to possess the monotonicity property.

2.3. **A Strong Stability Requirement.** An essential property of monotonic relaxation systems is that the solution vector remains bounded by the equilibrium value at all times. To avoid unphysical solutions and numerical oscillations, we want our numerical method to possess an analogous property.

Definition 3. *Consider a monotonic relaxation ODE with initial conditions* $\boldsymbol{V}^n$ *and equilibrium point* $\boldsymbol{V}^{\mathrm{eq}}$*. Let the numerical solution be given through some operator* $\mathcal{S}(\Delta t)$ *as*

$$\boldsymbol{V}^{n+1} = \mathcal{S}(\Delta t)\boldsymbol{V}^n. \tag{15}$$

The operator $\mathcal{S}$ *will be denoted as* **monotonically asymptotically stable** *if it satisfies the following properties.*

MA1: *The operator is* **consistent** *with the relaxation system to be solved, i.e. the local truncation error is of at least second order in* Δt*.*

MA2: *The solution is unconditionally* **bounded** *by the equilibrium value, i.e.*

$$\begin{aligned} V_i^{n+1} &\in (V_i^n, V_i^{\mathrm{eq}}) \qquad \textit{for} \quad V_i^n < V_i^{\mathrm{eq}}, \\ V_i^{n+1} &= V_i^n \qquad\qquad \textit{for} \quad V_i^n = V_i^{\mathrm{eq}}, \\ V_i^{n+1} &\in (V_i^{\mathrm{eq}}, V_i^n) \qquad \textit{for} \quad V_i^n > V_i^{\mathrm{eq}} \end{aligned} \tag{16}$$

for all $i \in \{1, \ldots, N\}$ *and for all* Δt*.*

Common explicit methods typically do not possess this form of stability. For instance, the Forward Euler method satisfies the property MA2 only conditionally, with a strong restriction on the time step:

$$\frac{\Delta t}{\epsilon} < \min_i \left(\frac{V_i^{\mathrm{eq}} - V_i^n}{S_i(\boldsymbol{V}^n)} \right). \tag{17}$$

Implicit methods may however possess such strong stability, as exemplified as follows.

Proposition 3. *The backward Euler method, defined by*

$$\boldsymbol{V}^{n+1} = \boldsymbol{V}^n + \frac{\Delta t}{\epsilon}\boldsymbol{S}(\boldsymbol{V}^{n+1}), \tag{18}$$

is monotonically asymptotically stable in the sense of Definition 3.

Proof. It is well known and easy to check that the backward Euler method is consistent; i.e. the property MA1 is satisfied. We now prove the property MA2 by showing that we otherwise get contradictions. First, we note that the backward Euler method preserves V_i^n if and only if $V_i^n = V_i^{\mathrm{eq}}$. We now consider the case $V_i^{\mathrm{eq}} > V_i^n$. Assume that the solution $\boldsymbol{V}^{n+1}$ of (18) satisfies

$$V_i^{n+1} < V_i^n. \tag{19}$$

From (12), we then have $S_i(\boldsymbol{V}^{n+1}) > 0$ which inserted into (18) yields $V_i^{n+1} > V_i^n$, in contradiction to (19).

Similarly, assume that the solution $\boldsymbol{V}^{n+1}$ of (18) satisfies

$$V_i^{n+1} > V_i^{\mathrm{eq}}. \tag{20}$$

From (12), we then have $S_i(\boldsymbol{V}^{n+1}) < 0$ which inserted into (18) yields $V_i^{n+1} < V_i^n$, in contradiction to (20).

The same steps will prove the remaining case $V_i^{\mathrm{eq}} < V_i^n$. □

Implicit methods generally require the solution of a system of nonlinear equations, which raises its own computational efficiency and robustness issues. This motivates the *explicit* monotonically asymptotically stable method presented in the following.

Definition 4. *The numerical method given by*

$$V_i^{n+1} = V_i^n + (V_i^{\mathrm{eq}} - V_i^n)\left(1 - \exp\left(-\frac{\Delta t}{\tau_i}\right)\right), \tag{21}$$

where

$$\tau_i = \epsilon \frac{V_i^{\mathrm{eq}} - V_i^n}{S_i(\boldsymbol{V}^n)}, \tag{22}$$

will be denoted as the **ASY1** *method.*

The ASY1 method may be straightforwardly derived by insisting that it should satisfy the following natural conditions:

(i) The numerical solution decays to the equilibrium value exponentially as the time step is increased;
(ii) The time constant of the exponential decay is chosen to make the method consistent to first order in Δt with the original ODE.

Now define

$$\delta_i = V_i^{\mathrm{eq}} - V_i^n. \tag{23}$$

It then follows from Proposition 1 that

$$\lim_{\delta_i \to 0}\left(\frac{\epsilon}{\tau_i}\right) = \frac{\partial S_i}{\partial V_i}, \tag{24}$$

hence (21) remains valid also for vanishing δ_i. However, to avoid numerical problems, one may consider replacing (21) with the limit (24) if δ_i becomes very small.

Proposition 4. *Let V_i^n be given by the ASY1 method of Definition 4. Then the local error*

$$\mathcal{E}_i^n = V_i^n - V_i(t^n) \tag{25}$$

satisfies the inequality

$$|\mathcal{E}_i(t)| \leq KC\delta\left(\frac{\Delta t}{\epsilon}\right)^2, \tag{26}$$

where

$$\delta = \max_j |\delta_j|, \tag{27}$$

$$C = \sup_{\boldsymbol{V}\in\mathcal{D}}\left|\frac{\partial^2 S}{\partial V_j \partial V_k}\right|, \tag{28}$$

and

$$K = \frac{\delta\left(N - \frac{1}{2}\right)}{|S_i(\boldsymbol{V}^n)|}\max_j\left(\frac{S_j(\boldsymbol{V}^n)}{\delta_j}\right)\left(|S_i(\boldsymbol{V}^n)| + \frac{1}{2}C\delta^2\left(N - \frac{1}{2}\right)\right). \tag{29}$$

Proof. From Lemma 9 and the variable transformations (95) and (96) in Appendix A, we directly obtain that K must satisfy

$$K \geq \frac{1}{2}\frac{\delta\left(N-\frac{1}{2}\right)}{|S_i\left(\boldsymbol{V}^n\right)|}\left(\frac{[S_i(\boldsymbol{V}^n)]^2}{|\delta_i|}+\max_j\left(\frac{S_j\left(\boldsymbol{V}^n\right)}{\delta_j}\right)\left(|S_i\left(\boldsymbol{V}^n\right)|+C\delta^2\left(N-\frac{1}{2}\right)\right)\right), \tag{30}$$

taking into account that the definition (25) scales the error (134) from the Appendix with a factor δ_i. The result then follows from further applying the inequality

$$\frac{S_i(\boldsymbol{V}^n)}{\delta_i} \leq \max_j\left(\frac{S_j\left(\boldsymbol{V}^n\right)}{\delta_j}\right). \tag{31}$$

□

Proposition 5. *The ASY1 method is monotonically asymptotically stable in the sense of Definition 3.*

Proof. It follows from Proposition 4 that the property MA1 is satisfied. From (12) and (22) it follows that the range of the exponential function is in the interval $(0, 1]$. Hence the property MA2 is satisfied. □

Remark 1. *Note that the ASY1 method* (21) *inserts a numerical "barrier" at the point $V_i = V_i^{\text{eq}}$ through which the solution can never pass. Hence the method cannot be consistent unless this barrier is also present in the underlying mathematical equation, as is the case for monotonic relaxation ODEs. This monotonicity property is explicitly needed for the error analysis in Appendix A.*

In some applications, the equilibrium value may be trivially calculated. For example, the relaxation term can represent some friction that drives a velocity to zero. In other cases, for instance flows involving phase transfers [7]*, equilibrium calculations can be computationally expensive or only approximately available. In such cases, underestimating the distance from the initial value to the equilibrium state leads to a loss of consistency of the ASY1 method. Overestimating this distance retains consistency, but makes the method behave more like the Forward Euler method.*

2.4. **Accuracy Near Equilibrium.** The exponential function employed in (21) is of course only one of many functions that asymptotically approaches a limit value. However, it becomes the natural choice as it corresponds to the *exact* solution for linear monotonic relaxation problems. We have the following proposition.

Proposition 6. *Let the ASY1 method of Definition 4 be applied to a monotonic relaxation ODE. Then the error $\mathcal{E}_i$ satisfies the inequality*

$$|\mathcal{E}_i| \leq \frac{C\delta^2\delta_i\left(N-\frac{1}{2}\right)}{S_i\left(\boldsymbol{V}^n\right)} \quad \forall \Delta t \geq 0. \tag{32}$$

Proof. The result follows from Lemma 6 in Appendix A with the definitions (102) and (123). Herein, it must be taken into account that the definition (25) scales the error (134) from the Appendix with a factor δ_i. □

We remark the following.

- For linear systems, we have $C = 0$ and hence $\mathcal{E}_i = 0$ for all $t \geq 0$.
- In general, S_0 can be arbitrarily close to zero. However, in the case that S_i decays *linearly* to zero at equilibrium, i.e.

$$L = \frac{\partial S_i}{\partial V_i} \neq 0 \quad \text{for} \quad V_i = V_i^{\text{eq}}, \tag{33}$$

then S_i will be of order δ_i and the error $\mathcal{E}_i$ will be of order δ^2 for sufficiently small δ.

In the case that S_i decays linearly, the error will decay *exponentially* with the time step, as described in the following.

Proposition 7. *Assume that $L \neq 0$ and that we are sufficiently close to the equilibrium so that*

$$\delta < \frac{\alpha L}{2C\left(N - \frac{1}{2}\right)} \tag{34}$$

for some $0 < \alpha \leq 1$. Then the error $\mathcal{E}_i(t)$ satisfies the inequality

$$|\mathcal{E}_i(t)| < |\delta_i| \exp\left(-(1-\alpha)L\frac{t}{\epsilon}\right). \tag{35}$$

Proof. The foundation of the proof is given in Appendix A. From Lemma 4 we obtain the bound

$$|\mathcal{E}_i(t)| \leq |\delta_i| \exp\left(-W\frac{t}{\epsilon}\right), \tag{36}$$

where

$$W = \frac{S_i\left(\boldsymbol{V}^n\right)}{\delta_i} - C\delta\left(N - \frac{1}{2}\right). \tag{37}$$

Now we have

$$|S_i\left(\boldsymbol{V}^n\right)| \geq L\,|\delta_i| - \frac{1}{2}C\delta_i^2, \tag{38}$$

giving

$$W \geq L - \frac{1}{2}C\,|\delta_i| - C\delta\left(N - \frac{1}{2}\right) > L(1-\alpha) \tag{39}$$

where we have used (34). Now the result follows from (36). □

Notably, the error decreases exponentially with the stiffness of the system for linearly decaying source terms. Hence the apparent "stiffness sensitivity" of the error bound (26) becomes limited as the time step is increased.

2.5. **Second-Order Accuracy.** A general explicit two-stage Runge–Kutta scheme for the ODE (4) can be written in the form

$$\boldsymbol{V}^* = \boldsymbol{V}^n + a\frac{\Delta t}{\epsilon}\boldsymbol{S}(\boldsymbol{V}^n) \tag{40}$$

$$\boldsymbol{V}^{n+1} = \boldsymbol{V}^n + \frac{\Delta t}{\epsilon}\left(b_1\boldsymbol{S}(\boldsymbol{V}^n) + b_2\boldsymbol{S}(\boldsymbol{V}^*)\right), \tag{41}$$

for second-order accuracy the parameters a, b_1 and b_2 must satisfy (see for instance [21, Ch. 8]):

$$b_1 + b_2 = 1, \qquad ab_2 = \frac{1}{2}. \tag{42}$$

In this section, we construct a second-order version of the ASY method through a similar two-stage application of (21).

Definition 5. *The numerical method given by*

$$V_i^* = V_i^n + (V_i^{\mathrm{eq}} - V_i^n)\left(1 - \exp\left(-a\frac{\Delta t}{\tau_i}\right)\right) \tag{43}$$

$$V_i^{n+1} = V_i^n + (V_i^{\mathrm{eq}} - V_i^n)\left(1 - b_1\exp\left(-\frac{\Delta t}{\tau_i}\right) - b_2\exp\left(-\frac{\Delta t}{\tau_i^*}\right)\right), \tag{44}$$

where

$$\tau_i = \epsilon \frac{V_i^{\mathrm{eq}} - V_i^n}{S_i(\boldsymbol{V}^n)}, \qquad \tau_i^* = \epsilon \frac{V_i^{\mathrm{eq}} - V_i^*}{S_i(\boldsymbol{V}^*)}, \tag{45}$$

and the parameters a, b_1 *and* b_2 *satisfy*

$$b_1 + b_2 = 1, \qquad ab_2 = \frac{1}{2}, \tag{46}$$

as well as

$$b_2 \in (0, 1], \tag{47}$$

will be denoted as the **ASY2** *method.*

Proposition 8. *When applied to a monotonic relaxation ODE, the ASY2 method is identical to the exact solution to second order in* Δt *in the Taylor expansion.*

Proof. Expanding τ_i^* we obtain

$$\frac{1}{\tau_i^*} = \frac{1}{\tau_i}\left(1 + a\Delta t\left(\frac{1}{\tau_i} + \frac{1}{S_i(\boldsymbol{V}^n)}\sum_{k=1}^{N}\frac{\partial S_i}{\partial V_k}(\boldsymbol{V}^n)\,\frac{S_k(\boldsymbol{V}^n)}{\epsilon}\right)\right) + \mathcal{O}(\Delta t^2), \tag{48}$$

where have used that

$$V_i^* = V_i^n + a\frac{\Delta t}{\epsilon}S_i(\boldsymbol{V}^n) + \mathcal{O}(\Delta t^2), \tag{49}$$

$$S_i(\boldsymbol{V}^*) = S_i(\boldsymbol{V}^n) + a\frac{\Delta t}{\epsilon}\sum_{k=1}^{N}\frac{\partial S_i}{\partial V_k}(\boldsymbol{V}^n)\,S_k(\boldsymbol{V}^n) + \mathcal{O}(\Delta t^2). \tag{50}$$

Substituting (48) into (44) and expanding the exponential function we obtain

$$\begin{aligned} V_i^{n+1} &= V_i^n + \frac{\Delta t}{\epsilon}S_i(\boldsymbol{V}^n)\,(b_1 + b_2) \\ &+ \frac{1}{2}\frac{\Delta t^2}{\epsilon^2}\left((2ab_2 - b_1 - b_2)\,\frac{S_i(\boldsymbol{V}^n)^2}{V_i^{\mathrm{eq}} - V_i^n} + 2ab_2\sum_{k=1}^{N}\frac{\partial S_i}{\partial V_k}(\boldsymbol{V}^n)\,S_k(\boldsymbol{V}^n)\right) + \mathcal{O}(\Delta t^3), \end{aligned} \tag{51}$$

whereas the exact solution satisfies

$$V_i(t^n + \Delta t) = V_i^n + \frac{\Delta t}{\epsilon}S_i(\boldsymbol{V}^n) + \frac{1}{2}\frac{\Delta t^2}{\epsilon^2}\sum_{k=1}^{N}\frac{\partial S_i}{\partial V_k}(\boldsymbol{V}^n)\,S_k(\boldsymbol{V}^n) + \mathcal{O}(\Delta t^3). \tag{52}$$

Now using (46) we may write

$$V_i(t^n + \Delta t) - V_i^{n+1} = \mathcal{O}(\Delta t^3) \quad \forall V_i^n \neq V_i^{\mathrm{eq}}. \tag{53}$$

We finally observe that the ASY2 method respects the limit

$$\lim_{V_i^n \to V_i^{\mathrm{eq}}} V_i^{n+1} = V_i^{\mathrm{eq}}. \tag{54}$$

□

Proposition 9. *The ASY2 method is monotonically asymptotically stable in the sense of Definition 3.*

Proof. The property MA1 follows immediately from Proposition 8. From (12), it follows that the exponential functions of (44) are bounded by the interval $(0, 1]$. The property MA2 then follows from (46)–(47). □

3. Hyperbolic Relaxation Systems

A hyperbolic relaxation system can be written in general quasilinear form as follows [26]:

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{A}(\boldsymbol{U})\frac{\partial \boldsymbol{U}}{\partial x} = \frac{1}{\epsilon}\boldsymbol{R}(\boldsymbol{U}), \tag{55}$$

where the matrix $\boldsymbol{A}$ is assumed to be diagonalizable with real eigenvalues in the domain of interest. In the context of (2), $\boldsymbol{A}$ is given by

$$\boldsymbol{A}(\boldsymbol{U}) = \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{U}}. \tag{56}$$

Such systems model many relevant physical problems, such as two-phase flows which are locally not in thermodynamic equilibrium [7, 8, 32, 37].

The limiting process $\epsilon \to 0$ in systems in the form (55) was extensively analysed by Liu [24] and Chen et al. [4], with a particular focus on the relationship between stability and wave propagation. It is of high interest to obtain good numerical methods for systems in the form (55) when the relaxation source term is stiff; i.e. the parameter ϵ is so small that the time scales associated with the homogeneous system:

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{A}(\boldsymbol{U})\frac{\partial \boldsymbol{U}}{\partial x} = 0 \tag{57}$$

are significantly larger than the time scales associated with the relaxation part:

$$\frac{\partial \boldsymbol{U}}{\partial t} = \frac{1}{\epsilon}\boldsymbol{R}(\boldsymbol{U}). \tag{58}$$

Several approaches have been proposed in the literature. These may be roughly divided into *splitting* and *unsplit* methods [29].

3.1. Numerical Methods.

We assume a uniform computational grid, and let $\boldsymbol{U}_j^n$ denote the cell averages of $\boldsymbol{U}$ in the cell $[x_{j-1/2}, x_{j+1/2}]$ at time t^n. Let $\mathcal{H}(t)$ be the operator that advances the system (57) forward in time, and let $\mathcal{S}(t)$ be the corresponding stiff ODE operator for the system (58). Then we may consider two main classes of splitting methods [18]:

- *Godunov splitting*:

$$\boldsymbol{U}^{n+1} = \mathcal{S}\left(\Delta t\right) \circ \mathcal{H}\left(\Delta t\right) \boldsymbol{U}^n, \tag{59}$$

- *Strang splitting* [34]:

$$\boldsymbol{U}^{n+1} = \mathcal{H}\left(\frac{1}{2}\Delta t\right) \circ \mathcal{S}\left(\Delta t\right) \circ \mathcal{H}\left(\frac{1}{2}\Delta t\right) \boldsymbol{U}^n. \tag{60}$$

Godunov splitting is first-order accurate, whereas Strang splitting is second-order accurate provided that both $\mathcal{H}$ and $\mathcal{S}$ are second-order accurate operators. In particular, Strang splitting applied to (57)–(58) is second-order accurate for any fixed ϵ and sufficiently small Δt. However, as emphasized by Pareschi and Russo [29], and proved by Jin [19], the method in general degenerates to first order in the limit $\epsilon \to 0$. Although this limit may never be fully realized in practical applications, this is nevertheless an undesirable property. Following the terminology of [29], we will denote schemes that retain their order of accuracy also in the limit $\epsilon \to 0$ as *asymptotically accurate*.

Jin [19] proposed an asymptotically second-order accurate splitting method based on two-stage Runge–Kutta time integration. This paved the way for a currently

popular class of methods; implicit-explicit (IMEX) Runge–Kutta methods [2, 3, 29] where an explicit discretization is applied to the flux terms and an implicit one to the source terms. This provides a general framework for achieving high-order asymptotic accuracy.

However, implicit methods involve the need to solve systems of nonlinear equations at each time step. Explicit methods do not suffer from this inconvenience, and would be preferable if applicable. In the context of (55), some properties of the ASY methods appear at first sight to be interesting. In particular:

- For stiff relaxation systems in the form (55), we may wish to employ a time step that is adapted to the hyperbolic dynamics (57). Such a time step may be excessively large for the relaxation part (58), and could potentially lead to an unphysical numerical solution, invalidating the simulation. For instance, if the relaxation process represents phase transitions, a too large time step could lead to one phase having a negative mass. The *Monotonic Asymptotic Stability* property (Propositions 5 and 9) would guarantee that this could never happen.
- Solutions to relaxation systems in the form (55) tend to remain close to an equilibrium state. This motivates a numerical method with a high degree of asymptotic accuracy (Propositions 6 and 7).

Nevertheless, to avoid order degeneracy in the stiff limit, we will have to overcome the challenge that the relaxation system (58) is intimately coupled to the hyperbolic part (57). This issue will not be explored in the current paper. In the next section, we will focus on demonstrating the beneficial properties of the ASY method when applied to a stand-alone monotonic relaxation ODE. We will then make some preliminary investigations on the applicability of the ASY method in the context of hyperbolic relaxation systems, by employing the simple splitting approach described above.

4. Numerical Examples

The aim of this section is to numerically illustrate the properties formally derived in Section 2, as well as getting some indications on the potential applicability of our methods to practical problems. To this end, we first construct a monotonic system of relaxation ODEs where the source term has a limited domain of definition. While the ASY methods guarantee that the solution will remain in this domain, some natural alternative methods here yield invalid solutions if the time step is chosen too large.

We then consider the *granular gas flow* model studied by Serna and Marquina [33], as this example allows for comparing the performance of our method to results existing in the literature. Throughout this section, we will use the parameter

$$a = 1 \tag{61}$$

for the ASY2 method of Definition 5. By this choice, we only need two evaluations of the exponential function in (43)–(44).

4.1. A Nonlinear Monotonic System.

In the context of (4), we consider the following expression for $\boldsymbol{S}(\boldsymbol{V})$:

$$\boldsymbol{V} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \qquad \boldsymbol{S}(\boldsymbol{V}) = -\begin{bmatrix} V_1\left(\sqrt{V_1} + V_2\right) \\ V_2\left(\sqrt{V_2} + V_1\right) \end{bmatrix}. \tag{62}$$

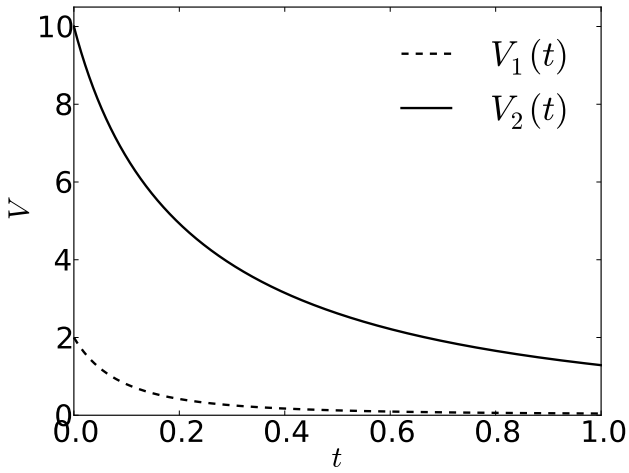


FIGURE 1. The time evolution of the solution vector $\boldsymbol{V}(t)$ for the benchmark relaxation ODE.

We may verify that this is a monotonic relaxation system in the sense of Definition 2, and that $\boldsymbol{S}(\boldsymbol{V})$ takes on real values only if $V_1, V_2 \geq 0$. The equilibrium value is

$$\boldsymbol{V}^{\text{eq}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{63}$$

For our numerical test, we will solve the ODE defined by (62) using $\epsilon = 1$ for

$$t \in [0, 1] \tag{64}$$

with the initial condition

$$\boldsymbol{V}(0) = \boldsymbol{V}_0 = \begin{bmatrix} 2 \\ 10 \end{bmatrix}. \tag{65}$$

The solution $\boldsymbol{V}(t)$ for $t \in [0, 1]$ is plotted in Figure 1, with the end state

$$\boldsymbol{V}(1) \approx \begin{bmatrix} 4.34664 \cdot 10^{-2} \\ 1.28793 \end{bmatrix}. \tag{66}$$

A phase diagram is shown in Figure 2. Herein, the orbit containing the point $\boldsymbol{V}_0$ is shown as a solid line.

4.1.1. *Numerical Methods.* In this numerical test, we will compare the ASY1 and ASY2 methods of Definitions 4 and 5 to some classical methods. In particular, we will consider the following Runge–Kutta methods:

RK1: The first order Forward Euler scheme:

$$\boldsymbol{V}^{n+1} = \boldsymbol{V}^n + \frac{\Delta t}{\epsilon} \boldsymbol{S}(\boldsymbol{V}^n). \tag{67}$$

RK2: The second order Heun's Method:

$$\boldsymbol{V}^* = \boldsymbol{V}^n + \frac{\Delta t}{\epsilon} \boldsymbol{S}(\boldsymbol{V}^n), \tag{68}$$

$$\boldsymbol{V}^{n+1} = \boldsymbol{V}^n + \frac{\Delta t}{2\epsilon} \left(\boldsymbol{S}(\boldsymbol{V}^n) + \boldsymbol{S}(\boldsymbol{V}^*) \right). \tag{69}$$

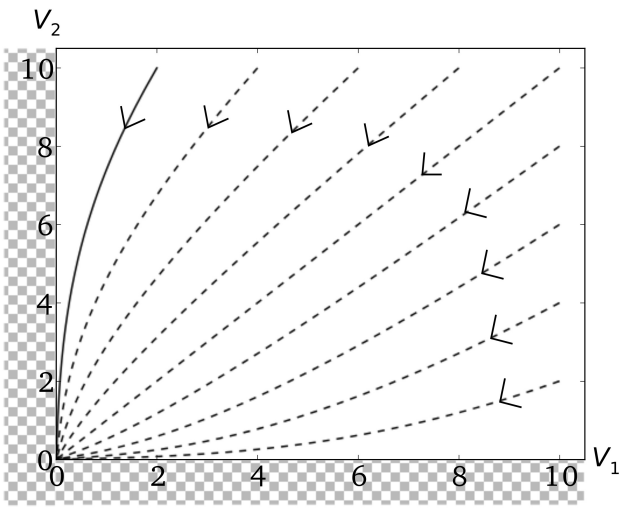


FIGURE 2. Phase diagram for the benchmark relaxation ODE. In the checkerboarded domain, the source term becomes complex.

For simplicity, we will employ the classical exponential integrator of Lawson [23], as described in Section 2.1. Herein, in the context of the splitting (7), we choose the linearization

$$\boldsymbol{L} = \boldsymbol{A}(\boldsymbol{V}_0), \tag{70}$$

where $\boldsymbol{A}$ is the Jacobian matrix

$$\boldsymbol{A}(\boldsymbol{V}) = \frac{1}{\epsilon}\frac{\mathrm{d}\boldsymbol{S}(\boldsymbol{V})}{\mathrm{d}\boldsymbol{V}}. \tag{71}$$

First and second-order versions of Lawson's method are defined as follows:

EXP1: We use the RK1 scheme defined above to solve for the variable $\boldsymbol{W}$.
EXP2: We use the RK2 scheme defined above to solve for the variable $\boldsymbol{W}$.

4.1.2. *Numerical Results.* A reference solution was calculated at $t = 1.0$ using the RK2 scheme with $\Delta t = 10^{-10}$, and the error $\mathcal{E}$ for the different schemes was calculated using the Euclidian norm at the point $t = 1.0$:

$$\mathcal{E} = \sqrt{\left(V_1^{\mathrm{ref}} - V_1^{\mathrm{num}}\right)^2 + \left(V_2^{\mathrm{ref}} - V_2^{\mathrm{num}}\right)^2}. \tag{72}$$

Table 1 shows the error for the different schemes for Δt ranging from 10^{-6} to 1.0.

The numerical order of convergence n was estimated using

$$n = \log_{10}\left(\frac{\mathcal{E}^{i+1}}{\mathcal{E}^i}\right) \tag{73}$$

where $\mathcal{E}^i$ denotes the error when using the step size $\Delta t = 10^i$. Table 2 shows the estimated order of convergence using the errors from Table 1.

TABLE 1. The error $\mathcal{E}$ at $t = 1.0$ for the numerical solution of the ODE (62) with $\boldsymbol{V}(t = 0) = [2, 10]^T$.

Δt	RK1	RK2	EXP1	EXP2	ASY1	ASY2
10^{-6}	$3.812{\cdot}10^{-7}$	$8.960{\cdot}10^{-13}$	$3.621{\cdot}10^{-7}$	$9.866{\cdot}10^{-14}$	$8.333{\cdot}10^{-8}$	$6.344{\cdot}10^{-14}$
10^{-5}	$3.812{\cdot}10^{-6}$	$8.960{\cdot}10^{-11}$	$3.621{\cdot}10^{-6}$	$9.869{\cdot}10^{-12}$	$8.333{\cdot}10^{-7}$	$6.346{\cdot}10^{-12}$
10^{-4}	$3.812{\cdot}10^{-5}$	$8.964{\cdot}10^{-9}$	$3.620{\cdot}10^{-5}$	$9.904{\cdot}10^{-10}$	$8.333{\cdot}10^{-6}$	$6.356{\cdot}10^{-10}$
10^{-3}	$3.813{\cdot}10^{-4}$	$9.005{\cdot}10^{-7}$	$3.614{\cdot}10^{-4}$	$1.025{\cdot}10^{-7}$	$8.327{\cdot}10^{-5}$	$6.458{\cdot}10^{-8}$
10^{-2}	$3.830{\cdot}10^{-3}$	$9.419{\cdot}10^{-5}$	$3.552{\cdot}10^{-3}$	$1.392{\cdot}10^{-5}$	$8.275{\cdot}10^{-4}$	$7.494{\cdot}10^{-6}$
10^{-1}	NaN	NaN	$3.233{\cdot}10^{-2}$	$5.858{\cdot}10^{-3}$	$7.929{\cdot}10^{-3}$	$1.908{\cdot}10^{-3}$
1	NaN	NaN	NaN	NaN	$4.344{\cdot}10^{-2}$	$8.964{\cdot}10^{-1}$

TABLE 2. The numerical order of convergence n at $t = 1.0$ for the solution of the ODE (62) with $\boldsymbol{V}(t = 0) = [2, 10]^T$.

Δt	RK1	RK2	EXP1	EXP2	ASY1	ASY2
10^{-6}	1.000	2.000	1.000	2.000	1.000	2.000
10^{-5}	1.000	2.000	1.000	2.002	1.000	2.001
10^{-4}	1.000	2.002	0.999	2.015	1.000	2.007
10^{-3}	1.002	2.020	0.993	2.133	0.997	2.065
10^{-2}	NaN	NaN	0.959	2.624	0.981	2.406
10^{-1}	NaN	NaN	NaN	NaN	0.739	2.672

4.1.3. *Interpretation of the Results.* We observe that both the RK and EXP schemes overshoot the equilibrium value for the largest time steps, producing complex numbers in the source term (62). This illustrates the situation that forms the primary motivation for the ASY methods. As stated by Propositions 5 and 9, the ASY methods yield physically valid solutions with no restrictions on the time step.

We observe that all methods display the expected numerical order of convergence. For this test case, the ASY methods consistently perform better than their RK and EXP counterparts for a given time step size.

The performance of the EXP methods could probably be improved by choosing a more representative integrator than the simple Lawson's method. On the other hand, the ASY methods still have the benefit of not depending on a splitting (7); they depend only on the equilibrium state $\boldsymbol{V}^{\text{eq}}$. Also, they do not require the calculation of any matrix exponential.

4.2. **A Granular-Gas Flow Model.** Granular gases have lately been the subject of considerable theoretical, numerical and experimental studies [9, 30, 29, 33, 31]. In this work we consider a continuum model for granular-gas flow, in which the dynamics are accounted for by a hyperbolic conservation law with relaxation. Our main motivation for choosing this example is the existence of previously published numerical results [20, 29, 33], to which our simulations may be compared.

In addition, the ASY methods should be well suited to the following features of the model:

- The relaxation part of the system is a monotonic nonlinear relaxation ODE.

- The equilibrium state corresponds to a granular temperature $T = 0$ and is hence easy to calculate.
- Numerically overshooting the equilibrium would be undesirable, as it would lead to the unphysical state $T < 0$, yielding complex values in the source term.

4.2.1. *Fluid-Mechanical Equations.* The dynamics of a one-dimensional granular-gas flow under the influence of gravity, in the form considered by Serna and Marquina [33], can be described by the Euler-like equations

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} = 0, \tag{74a}$$

$$\frac{\partial (\rho u)}{\partial t} + \frac{\partial (\rho u^2 + p)}{\partial x} = \rho g, \tag{74b}$$

$$\frac{\partial E}{\partial t} + \frac{\partial\, u(E + p)}{\partial x} = \Theta + \rho g u. \tag{74c}$$

In the above, ρ is the density, u is the velocity, p is the pressure, g is the gravitational acceleration, E is the energy density and Θ is the rate of energy loss due to inelastic collisions. The energy density consists of both kinetic and internal energy and is given by $E = (1/2)\rho u^2 + (3/2)\rho T$, where T is the granular temperature.

Following Serna and Marquina [33], we use an energy-loss term based on Haff's cooling law [10], given by

$$\Theta(\rho, T) = -\frac{12}{\sqrt{\pi}} \frac{1 - e^2}{\sigma} \rho T^{3/2} G(\nu), \tag{75}$$

where σ is the particle diameter and $e \in [0, 1]$ is the restitution coefficient. For $e = 1$ we recover a fully elastic model. The statistical correlation function $G(\nu)$ is given by

$$G(\nu) = \nu \left(1 - \left(\frac{\nu}{\nu_M} \right)^{\frac{3}{4}\nu_M} \right)^{-1}, \tag{76}$$

where $\nu = (\pi/6)\rho\sigma^3$ is the volume fraction and ν_M is the maximal volume fraction.

The pressure is determined by a granular equation of state (EOS), introduced by Goldshtein and Shapiro [9], given by

$$p(\rho, T) = T\rho(1 + 2(1 + e)G(\nu)). \tag{77}$$

4.2.2. *The Relaxation ODE.* Within the splitting (57)–(58), we obtain

$$\boldsymbol{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix} \qquad \text{and} \qquad \frac{1}{\epsilon}\boldsymbol{R}(\boldsymbol{U}) = \begin{bmatrix} 0 \\ 0 \\ \Theta(\rho, T) \end{bmatrix}. \tag{78}$$

For any initial condition

$$\boldsymbol{U}_0 = \begin{bmatrix} \rho_0 \\ \rho_0 u_0 \\ E_0 \end{bmatrix}, \tag{79}$$

this may be written in the reduced form (4) with

$$V(\boldsymbol{U}) = E, \tag{80}$$

$$\frac{1}{\epsilon} S(V) = -8\sqrt{\frac{2}{3\pi\rho_0}} \frac{1 - e^2}{\sigma} \left(V - \frac{1}{2}\rho_0 u_0^2 \right)^{3/2} G(\nu_0). \tag{81}$$

Furthermore, for any V we can reconstruct the full state vector $\boldsymbol{U}$ as

$$\boldsymbol{U}(V) = \begin{bmatrix} \rho_0 \\ \rho_0 u_0 \\ V \end{bmatrix}. \tag{82}$$

Note that the ODE defined by (81) can be integrated analytically to yield

$$V(t) = \frac{1}{2}\rho_0 u_0^2 + \left(E_0 - \frac{1}{2}\rho_0 u_0^2\right)\left(1 + 4t\sqrt{\frac{2}{3\pi\rho_0}}\frac{1-e^2}{\sigma}G(\nu_0)\sqrt{E_0 - \frac{1}{2}\rho_0 u_0^2}\right)^{-2}. \tag{83}$$

As the purpose of this section is to illustrate the use of the ASY methods in the context of hyperbolic conservation laws with relaxation, we will integrate (81) numerically rather than make use of this analytical expression.

4.2.3. *Numerical Method.* In order to numerically demonstrate the ASY methods on the granular-gas model described in Section 4.2, we use a fractional-step method as described in Section 3.1. This means that we need a numerical solver for the hyperbolic part (57) to use in tandem with the ODE solver; we will use the MUSTA method of Toro [35], augmented with the MUSCL approach of van Leer [36].

We consider a uniform grid in space and time, and denote $t^n = t_0 + n\,\Delta t$ and $x_j = x_0 + j\,\Delta x$. For a first-order accurate numerical scheme, we advance the solution $\boldsymbol{U}_j^n$ forward in time by using

$$\boldsymbol{U}_j^{n+1} = \boldsymbol{U}_j^n + \mathcal{F}_j^n \Delta t, \tag{84}$$

where

$$\mathcal{F}_j^n = \frac{1}{\Delta x}\left(\boldsymbol{F}_{j-1/2}^n - \boldsymbol{F}_{j+1/2}^n\right) + \boldsymbol{\mathcal{Q}}(\boldsymbol{U}_j^n). \tag{85}$$

In the above, $\boldsymbol{F}_{j+1/2}^n$ is the numerical approximation to the inter-cell flux and $\boldsymbol{\mathcal{Q}}(\boldsymbol{U}_j^n)$ are local source terms other than relaxation terms. For the granular-gas model, $\boldsymbol{\mathcal{Q}}(\boldsymbol{U})$ will be the gravity source terms.

In the Multi-Stage (MUSTA) approach, the inter-cell flux is calculated by solving the local Riemann problem at each cell interface on a local grid [35]. The solution on the local grid is then advanced in several stages giving an approximation to the inter-cell flux. In our application, we will use four local grid cells and two local iteration steps. The CFL number of the local grid is kept the same as on the global grid.

4.2.4. *High Resolution.* In a high resolution (second order) extension to the MUSTA scheme, we employ a second-order strong-stability-preserving (SSP) Runge–Kutta method to advance the solution forward in time. The two-stage scheme is given by

$$\begin{aligned} \boldsymbol{U}_j^* &= \boldsymbol{U}_j^n + \mathcal{F}_j^n \Delta t, \\ \boldsymbol{U}_j^{n+1} &= \frac{1}{2}\boldsymbol{U}_j^n + \frac{1}{2}\boldsymbol{U}_j^* + \frac{1}{2}\mathcal{F}_j^* \Delta t. \end{aligned} \tag{86}$$

In order to obtain second-order accuracy in space, a piecewise linear MUSCL interpolation [28, 36] was used. For the granular-gas model, the variables used in the interpolation were given by

$$\boldsymbol{W} = \begin{bmatrix} \rho & v & p \end{bmatrix}^T. \tag{87}$$

We reconstruct these variables to the right and to the left of the cell interface as

$$\boldsymbol{W}^{\mathrm{R}}_{j+1/2} = \boldsymbol{W}_{j+1} - \frac{\Delta x}{2}\boldsymbol{\sigma}_{j+1} \quad \text{and} \quad \boldsymbol{W}^{\mathrm{L}}_{j+1/2} = \boldsymbol{W}_{j} + \frac{\Delta x}{2}\boldsymbol{\sigma}_{j}, \tag{88}$$

respectively. The cell slopes $\boldsymbol{\sigma}_j$ are calculated using a *minmod* slope, given by

$$\sigma_{j,i} = \mathrm{minmod}\left(\frac{W_{j,i} - W_{j-1,i}}{\Delta x}, \frac{W_{j+1,i} - W_{j,i}}{\Delta x}\right), \tag{89}$$

where the minmod function is defined as

$$\mathrm{minmod}(a,b) = \begin{cases} 0 & \text{if } ab \le 0 \\ a & \text{if } |a| < |b| \text{ and } ab > 0 \\ b & \text{if } |b| < |a| \text{ and } ab > 0 \end{cases}. \tag{90}$$

The reconstructed values at the interface are then used for the Riemann problem on the local MUSTA grid, in order to obtain second-order accuracy in space. We refer to the high-resolution scheme as MUSCL-MUSTA.

4.2.5. *Test Case: Granular-Gas Tube.* In this section we use the ASY integrators as a part of a fractional-step method in order to compare with previously reported results for the granular-gas model.

We consider the case of a granular gas in a vertical tube hitting a solid wall at the bottom end, as used by Serna and Marquina [33] and also Pareschi and Russo [29]. A highly similar simulation was presented by Kamath and Du [20].

The 0.1 m tube is initially filled with a granular gas with $\nu = 0.018\,\mathrm{kg}$, velocity $0.18\,\mathrm{m/s}$ and pressure $p = 1589.26\,\mathrm{Pa}$. We use the gravitational acceleration $g = 9.8\,\mathrm{m/s}$, the restitution coefficient $e = 0.97$, $\nu_M = 0.65\,\mathrm{kg}$ and the particle diameter $\sigma = 10^{-3}\,\mathrm{m}$. The left boundary condition is given by an incoming flow consistent with the initial condition. At the right end of the domain we used a reflective boundary condition.

Simulations were carried out using 200 computational cells and a CFL number of 0.4. Figure 3 shows the results for the packing fraction, granular temperature and pressure at $t = 0.23\,\mathrm{s}$, using the MUSTA-ASY1 scheme with Godunov splitting and the MUSCL-MUSTA-ASY2 scheme with Strang splitting. The reference solution was computed using the MUSCL-MUSTA-ASY2 scheme with 10 000 cells.

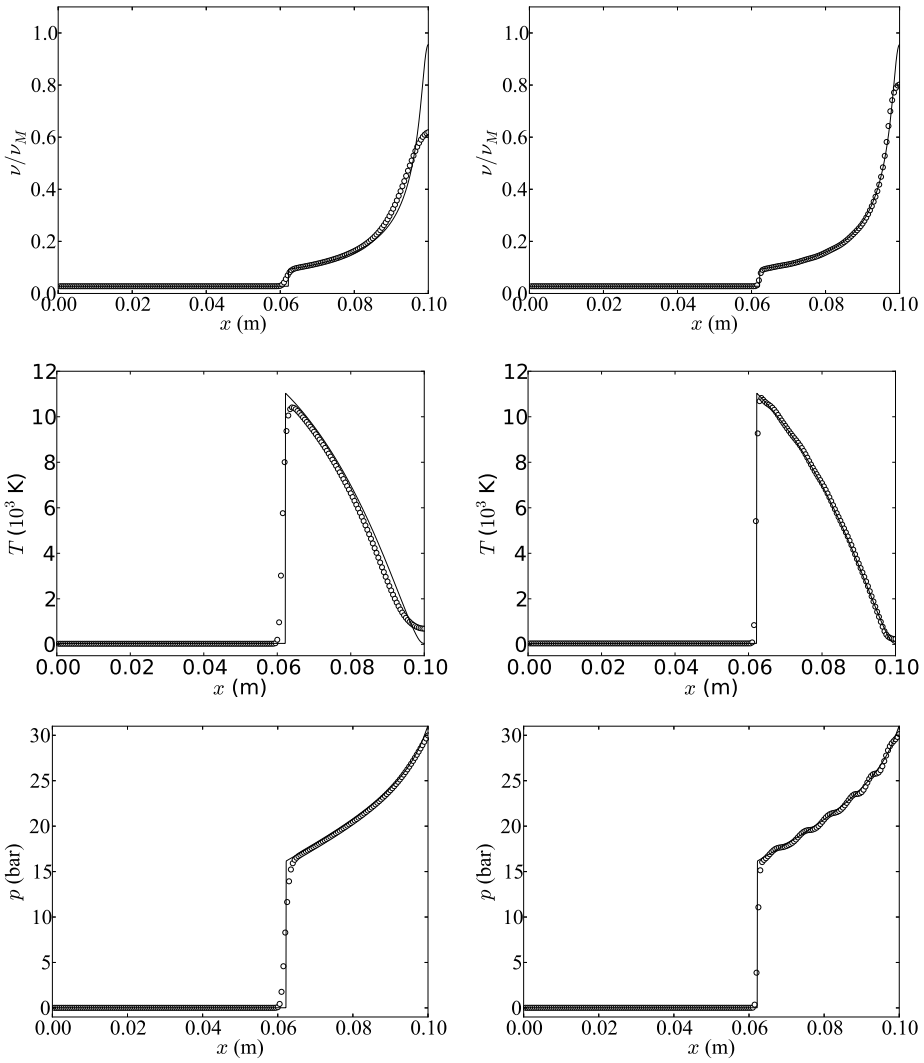
The results show a shock being formed when the gas hits the solid wall. The shock propagates backwards and the gas continues to compress against the wall until the maximum volume fraction is reached at the right boundary. It is also at the right boundary the difference between the first and second-order schemes is most prominent.

4.2.6. *Test Case: Stiffened Granular-Gas Tube.* In order to demonstrate the performance of the proposed numerical schemes in a stiff case, consider the artificially scaled relaxation term

$$\Theta_\varepsilon(\rho, T) = \frac{1}{\varepsilon}\Theta(\rho, T), \tag{91}$$

where $\varepsilon > 0$ is a stiffness parameter. This can be seen as a scaling of the particle diameter by a factor ε.

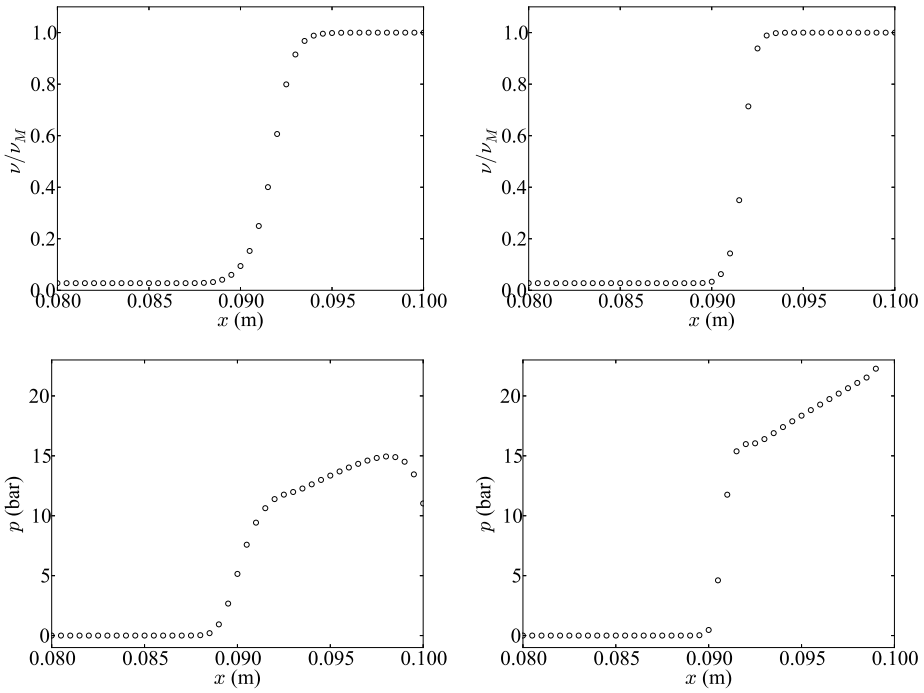
The numerical solution of the stiffened granular-gas tube was calculated using the same initial data, boundary conditions and numerical parameters as those used in Section 4.2.5. The stiffness parameter was $\varepsilon = 0.02$. Note that in this case, the problem is stiff in the sense that an explicit first-order Runge–Kutta step in



(a) MUSTA-ASY1 (b) MUSCL-MUSTA-ASY2

FIGURE 3. Granular-gas shock case at $t = 0.23\,\mathrm{s}$ for the MUSTA-ASY1 scheme and the MUSCL-MUSTA-ASY2 scheme. The solid line is the reference solution.

the ODE-part will overshoot equilibrium and produce an unphysical state. Figure 4 shows the packing fraction and pressure at $t = 0.23$ for the MUSTA-ASY1 and MUSCL-MUSTA-ASY2 schemes. We observe that, in contrast to the case from Section 4.2.5, the stiffened case reaches the maximal packing fraction in the right part of the tube. In this limit the speed of sound tends to infinity [33], which causes severe time-step restrictions. A reference solution was therefore not calculated in this case.



(a) MUSTA-ASY1 (b) MUSCL-MUSTA-ASY2

FIGURE 4. The stiffened granular-gas shock case at $t = 0.23\,\mathrm{s}$ for the MUSTA-ASY1 scheme and the MUSCL-MUSTA-ASY2 scheme.

4.2.7. *Interpretation of the Results.* The parameters used for the first granular-gas case do not lead to a high degree of stiffness. Hence this test case does not directly illustrate the asymptotic accuracy and robustness properties of the ASY methods. On the other hand, our choice of parameters allows us to compare our simulations to results previously reported in the literature [20, 29, 33]. In this respect, our results do not compare unfavourably in terms of accuracy and numerical robustness.

The artificially stiffened granular-gas case demonstrated the ability for the ASY1 and ASY2 methods to handle stiff source terms. The results show that the methods are stable even when reaching the close-packed limit.

These results, together with the properties formally derived in Section 2 and illustrated numerically in Section 4.1, indicate that the ASY methods show potential for being useful in the context of hyperbolic relaxation systems.

5. Summary

We have investigated a technique, based on exponential integration, for solving monotonic relaxation ODEs. First and second-order versions of our methods have been presented. We have proved that the resulting methods possess desirable accuracy and stability properties. In particular, for first-order corrections to the equilibrium value, the methods yield the exact solution.

Furthermore, the methods yield numerical solutions that are unconditionally bounded by the equilibrium state. This forms the main motivation behind our approach, and ensures a high degree of robustness in that unphysical solutions can be avoided.

These accuracy and robustness properties have been numerically demonstrated on a simple system of relaxation ODEs. We have argued that these properties may be particularly relevant in the context of hyperbolic conservation laws with relaxation. Through operator splitting, we have tested the viability of our approach to a model representing flow of granular gases with encouraging results.

The methods are, by design, applicable only to a restricted class of ODEs denoted as *monotonic relaxation ODEs*. This is both the weakness and strength of our methods. The strength lies in the fact that when the underlying equation system is monotonic, our methods will mimic its asymptotic behaviour in a simple and accurate manner. For monotonic systems, our methods do not require any calculation of the matrix exponential, and are uniquely determined by the equilibrium state of the given ODE.

Further work is needed to derive higher-order conditions for general multi-stage versions of the method. In the context of hyperbolic relaxation systems, it would be of high interest to systematically investigate unsplit versions of the approach, to avoid the order degeneracy in the stiff limit associated with operator splitting. Herein, ideas presented by Jin [19] may provide a useful starting point.

Appendix A. Error Analysis – Technical Details

A.1. **A Transformation of Variables.** To simplify the analysis, it will be convenient to write (1) in dimensionless form. We write

$$V_i(t) = \beta_i(t)V_i(0) + (1 - \beta_i(t))V_i^{\mathrm{eq}}, \tag{92}$$

where

$$\beta_i(0) = 1. \tag{93}$$

Now it follows from monotonicity that

$$\beta_i \in [0, 1], \tag{94}$$

and we define

$$\delta_i = -\frac{\mathrm{d}V_i}{\mathrm{d}\beta_i} = V_i^{\mathrm{eq}} - V_i(0). \tag{95}$$

We also introduce the rescaled time variable

$$\vartheta_i = \frac{S_i(0)t}{\epsilon\delta_i}, \tag{96}$$

giving

$$\mathrm{d}t = \frac{\epsilon\delta_i}{S_i(0)}\,\mathrm{d}\vartheta_i, \tag{97}$$

where we have used the shorthand

$$S_i(0) = S_i(\boldsymbol{V}(\boldsymbol{\beta}(0))). \tag{98}$$

Hence (1) can be written as

$$\frac{\mathrm{d}\beta_i}{\mathrm{d}\vartheta_i} = -\frac{S_i(\boldsymbol{V}(\boldsymbol{\beta}))}{S_i(0)} = -\xi_i(\boldsymbol{\beta}). \tag{99}$$

For a monotonic relaxation system, we have

$$\xi_i > 0 \quad \text{for} \quad \beta_k \in (0,1], \quad k \in \{1,\dots,N\}. \tag{100}$$

Also note that

$$\left|\frac{\partial^2 \xi_i}{\partial\beta_j\partial\beta_k}\right| = \left|\frac{\delta_j\delta_k}{S_i(0)}\right| \left|\frac{\partial^2 S_i}{\partial V_j \partial V_k}\right| \quad \forall i,j,k. \tag{101}$$

It will prove convenient to introduce the dimensionless constants

$$\mu_i = \max_j \left|\frac{\partial^2 \xi_i}{\partial\beta_i\partial\beta_j}\right|, \tag{102}$$

where the maximum is taken over all possible values of $\boldsymbol{\beta} \in [0,1]^N$.

Lemma 1. *The source term $\xi_i(\boldsymbol{\beta})$ satisfies the sharp inequality*

$$\xi_i(\boldsymbol{\beta}) \le \hat{\zeta}_i(\boldsymbol{\beta}), \tag{103}$$

where

$$\hat{\zeta}_i(\boldsymbol{\beta}) = \beta_i \left(1 - \frac{1}{2}\mu_i(1-\beta_i) + \mu_i \sum_{j=1}^{N}(1-\beta_j)\right). \tag{104}$$

Proof. Write ξ_i as a linear interpolant

$$\xi_i(\boldsymbol{\beta}^*(s)) = (1-s)\xi_i(\boldsymbol{\beta}^*(0)) + s\xi_i(\boldsymbol{\beta}^*(1)) + \mathcal{R}(s), \tag{105}$$

where

$$\beta_j^*(s) = \begin{cases} s & \text{if} \quad i = j \\ \beta_j & \text{otherwise.} \end{cases} \tag{106}$$

Then it follows from the error formula for polynomial interpolation that

$$|\mathcal{R}(\beta_i)| \le \frac{1}{2}\mu_i\beta_i(1-\beta_i), \tag{107}$$

giving

$$\xi_i(\boldsymbol{\beta}) \le \beta_i \left(\xi_i(\boldsymbol{\beta}^*(1)) + \frac{1}{2}\mu_i(1-\beta_i)\right). \tag{108}$$

Furthermore, we have

$$\xi_i(\boldsymbol{\beta}^*(1)) = 1 + \sum_{j\ne i} \frac{\partial\xi_i}{\partial\beta_j}(\bar{\boldsymbol{\beta}})(\beta_j - 1), \tag{109}$$

for some $\bar{\boldsymbol{\beta}} \in [0,1]^N$. It now follows from (13) that

$$\left|\frac{\partial\xi_i}{\partial\beta_j}(\bar{\boldsymbol{\beta}})\right| \le 0 + \mu_i\beta_i \le \mu_i. \tag{110}$$

Hence

$$\xi_i(\boldsymbol{\beta}^*(1)) \le 1 + \sum_{j\ne i}\left|\frac{\partial\xi_i}{\partial\beta_j}(\bar{\boldsymbol{\beta}})\right|(1-\beta_j) \le 1 + \mu_i\sum_{j\ne i}(1-\beta_j), \tag{111}$$

and the result follows from (108). Note that the inequality becomes an equality if all the second derivatives are equal, constant and positive, proving that the bound is sharp. □

Lemma 2. *The source term $\xi_i(\boldsymbol{\beta})$ satisfies the inequality*

$$\xi_i(\boldsymbol{\beta}) \geq \check{\zeta}_i(\boldsymbol{\beta}), \tag{112}$$

where

$$\check{\zeta}_i(\boldsymbol{\beta}) = \beta_i \left(1 + \frac{1}{2}\mu_i(1-\beta_i) - \mu_i \sum_{j=1}^{N}(1-\beta_j)\right). \tag{113}$$

Proof. The result follows from arguments fully analogous to the proof of Lemma 1. Note that the inequality becomes an equality if all the second derivatives are equal, constant and positive. However, unless

$$\mu_i \left(N - \frac{1}{2}\right) \leq 1 \tag{114}$$

$\check{\zeta}_i(\boldsymbol{\beta})$ will become negative at some points. Hence (114) must be satisfied for the bound to be sharp. □

Lemma 3. *Assume that the source term satisfies*

$$\check{\xi}_i(\boldsymbol{\beta}) \leq \xi_i(\boldsymbol{\beta}) \leq \hat{\xi}_i(\boldsymbol{\beta}) \quad \forall \boldsymbol{\beta} \in [0,1]^N. \tag{115}$$

Let $\hat{\beta}_i$ and $\check{\beta}_i$ be the solutions to the modified equations

$$\frac{\mathrm{d}\hat{\beta}_i}{\mathrm{d}\vartheta_i} = -\hat{\xi}_i(\hat{\boldsymbol{\beta}}), \qquad \hat{\beta}_i(0) = 1, \tag{116}$$

$$\frac{\mathrm{d}\check{\beta}_i}{\mathrm{d}\vartheta_i} = -\check{\xi}_i(\check{\boldsymbol{\beta}}), \qquad \check{\beta}_i(0) = 1. \tag{117}$$

Then

$$\hat{\beta}_i(\vartheta_i) \leq \beta_i(\vartheta_i) \leq \check{\beta}_i(\vartheta_i) \quad \forall \vartheta_i \geq 0. \tag{118}$$

Proof. Note that at $\vartheta_i = 0$ we have $\check{\beta}_i = \hat{\beta}_i = \beta_i = 1$. Hence, if there is some $\hat{\vartheta}_i > 0$ where $\hat{\beta}_i(\hat{\vartheta}_i) > \beta_i(\hat{\vartheta}_i)$ there must be some $\vartheta_i < \hat{\vartheta}_i$ where

$$\beta(\vartheta_i) = \hat{\beta}_i(\vartheta_i), \tag{119}$$

$$\frac{\mathrm{d}}{\mathrm{d}\vartheta_i}\left(\hat{\beta}_i - \beta_i\right) = \xi_i - \hat{\xi}_i > 0, \tag{120}$$

in contradiction to (115). Similarly, if there is some $\check{\vartheta}_i > 0$ where $\check{\beta}_i(\check{\vartheta}_i) < \beta_i(\check{\vartheta}_i)$ there must be some $\vartheta_i < \check{\vartheta}_i$ where

$$\beta(\vartheta_i) = \check{\beta}_i(\vartheta_i), \tag{121}$$

$$\frac{\mathrm{d}}{\mathrm{d}\vartheta_i}\left(\check{\beta}_i - \beta_i\right) = \xi_i - \check{\xi}_i < 0, \tag{122}$$

in contradiction to (115). □

Defining the shorthand

$$\eta_i = \mu_i \left(N - \frac{1}{2}\right), \tag{123}$$

we may now state the following result.

Lemma 4. *The solution $\beta_i(\vartheta_i)$ satisfies the inequality*

$$e^{-(1+\eta_i)\vartheta_i} \leq \beta_i(\vartheta_i) \leq e^{-(1-\eta_i)\vartheta_i}. \tag{124}$$

Proof. Define

$$\hat{\xi}_i(\boldsymbol{\beta}) = \beta_i\,(1+\eta_i) \geq \hat{\zeta}_i, \tag{125}$$

$$\check{\xi}_i(\boldsymbol{\beta}) = \beta_i\,(1-\eta_i) \leq \check{\zeta}_i, \tag{126}$$

and the result follows from Lemmas 1, 2 and 3. □

Now defining

$$\eta = \max_i \eta_i, \tag{127}$$

$$z_j = e^{-(1+\eta)\vartheta_j}, \tag{128}$$

$$\hat{z} = \min_j (z_j)\,, \tag{129}$$

we obtain the following lemma.

Lemma 5. *The source term $\xi_i(\boldsymbol{\beta})$ satisfies the inequality*

$$\beta_i\,(1-(1-\hat{z})\eta_i) \leq \xi_i \leq \beta_i\,(1+(1-\hat{z})\eta_i)\,. \tag{130}$$

Proof. From Lemmas 1, 2 and 4 we obtain

$$\beta_i\left(1-\frac{1}{2}\mu_i(1-\beta_i)-\mu_i(N-1)(1-\hat{z})\right) \leq \xi_i \leq \beta_i\left(1+\frac{1}{2}\mu_i(1-\beta_i)+\mu_i(N-1)(1-\hat{z})\right), \tag{131}$$

from which the result follows from (123). □

A.2. **The ASY Method.** In this dimensionless formulation, the ASY1 method is given as the exact solution $\tilde{\beta}_i$ to (99) with

$$\tilde{\xi}_i = \tilde{\beta}_i, \qquad \tilde{\beta}_i(0) = 1, \tag{132}$$

i. e.

$$\tilde{\beta}_i(\vartheta_i) = e^{-\vartheta_i}. \tag{133}$$

We now define the local error

$$E_i(\vartheta_i) = \tilde{\beta}_i(\vartheta_i) - \beta_i(\vartheta_i). \tag{134}$$

Lemma 6. *The error $E_i(\vartheta_i)$ satisfies the inequality*

$$|E_i(\vartheta_i)| \leq \eta_i \quad \forall \vartheta_i \geq 0. \tag{135}$$

Proof. Assume that the error is positive. It then follows from (124) that

$$E_i(\vartheta_i) \leq e^{-\vartheta_i}\left(1-e^{-\eta_i\vartheta_i}\right) \leq \left((1+\eta_i)^{-\frac{1+\eta_i}{\eta_i}}\right)\eta_i \leq \eta_i. \tag{136}$$

Assume that the error is negative and that $\eta < 1$. It then follows from (124) that

$$|E_i(\vartheta_i)| \leq e^{-\vartheta_i}\left(e^{\eta_i\vartheta_i}-1\right) \leq \left((1-\eta_i)^{\frac{1-\eta_i}{\eta_i}}\right)\eta_i \leq \eta_i. \tag{137}$$

To complete the proof, assume that the error is negative and that $\eta_i \geq 1$. Given $\beta_i \in [0,1]$, it then follows directly from the definition (134) that

$$|E_i(\vartheta_i)| \leq 1 \leq \eta_i. \tag{138}$$

□

A.3. **Temporal Accuracy.** We now define the parameter

$$r_i = \frac{1}{\vartheta_i} \max_j \vartheta_j \geq 1. \tag{139}$$

Lemma 7. *Assume that the source term is given by*

$$\xi_i(\vartheta_i) = \beta_i \left(1 + (1 - \hat{z})\eta_i\right). \tag{140}$$

Then the error satisfies

$$|E_i(\vartheta_i)| \leq \frac{1}{2}\eta_i \left(1 + r_i(1+\eta)\right)\vartheta_i^2 \quad \forall \vartheta_i \geq 0. \tag{141}$$

Proof. Observe that we have

$$E_i(0) = 0, \tag{142}$$
$$E_i'(0) = 0, \tag{143}$$
$$E_i''(\vartheta_i) = E_i(\vartheta_i) + \beta_i\eta_i \left(2(\hat{z}-1) - \eta_i(\hat{z}-1)^2 + \hat{z}r_i(1+\eta)\right), \tag{144}$$

where $E_i(\vartheta_i) \geq 0$. From Lemma 6 we find that

$$|E_i''(\vartheta_i)| \leq \eta_i \left(1 + r_i(1+\eta)\right), \tag{145}$$

from which the result follows. □

Lemma 8. *Assume that the source term is given by*

$$\xi_i(\vartheta_i) = \beta_i \left(1 - (1 - \hat{z})\eta_i\right). \tag{146}$$

Then the error satisfies

$$|E_i(\vartheta_i)| \leq \frac{1}{2}\eta_i \left(1 + r_i(1+\eta)\right)\vartheta_i^2 \quad \forall \vartheta_i \geq 0. \tag{147}$$

Proof. Observe that we have

$$E_i(0) = 0, \tag{148}$$
$$E_i'(0) = 0, \tag{149}$$
$$E_i''(\vartheta_i) = E_i(\vartheta_i) - \beta_i\eta_i \left(2(\hat{z}-1) + \eta_i(\hat{z}-1)^2 + \hat{z}r_i(1+\eta)\right), \tag{150}$$

where $E_i(\vartheta_i) \leq 0$. From Lemma 6 we find that

$$|E_i''(\vartheta_i)| \leq \eta_i \left(1 + r_i(1+\eta)\right), \tag{151}$$

from which the result follows. □

Lemma 9. *For any valid source term $\xi_i(\vartheta_i)$, the error $E_i(\vartheta_i)$ satisfies the inequality*

$$|E_i(\vartheta_i)| \leq \frac{1}{2}\eta_i \left(1 + r_i(1+\eta)\right)\vartheta_i^2 \quad \forall \vartheta_i \geq 0. \tag{152}$$

Proof. The result follows from Lemmas 3, 5, 7 and 8. □

Acknowledgements

The work of the second author was supported by A/S Norske Shell. The remaining authors were financed in part through the CO_2 Dynamics project. These authors acknowledge the support from the Research Council of Norway (189978), Gassco AS, Statoil Petroleum AS and Vattenfall AB. The work of the third author was supported in part by SINTEF Materials and Chemistry.

We are grateful to the anonymous reviewers for their very thorough reading of the manuscript. Their comments led to significant improvements to the original version of this paper.

References

[1] H. Berland, B. Owren and B. Skaflestad, *B*-series and order conditions for exponential integrators, *SIAM J. Numer. Anal.* **43**, 1715–1727, (2005).

[2] S. Boscarino, Error analysis of IMEX Runge–Kutta methods derived from differential-algebraic systems, *SIAM J. Numer. Anal.* **45**, 1600–1621, (2007).

[3] S. Boscarino and G. Russo, On a class of uniformly accurate IMEX Runge–Kutta schemes and applications to hyperbolic systems with relaxation, *SIAM J. Sci. Comput.* **31**, 1926–1945, (2009).

[4] G.-Q. Chen, C. D. Levermore and T.-P. Liu, Hyperbolic conservation laws with stiff relaxation terms and entropy, *Comm. Pure Appl. Math.* **47**, 787–830, (1994).

[5] S. M. Cox and P. C. Matthews, Exponential time differencing for stiff systems, *J. Comput. Phys.* **176**, 430–455, (2002).

[6] B. H. Ehle and J. D. Lawson, Generalized Runge–Kutta processes for stiff initial-value problems, *J. Inst. Maths. Applics* **16**, 11–21, (1975).

[7] T. Flåtten and H. Lund, Relaxation two-phase flow models and the subcharacteristic condition, *Math. Mod. Meth. Appl. S.* **21**, 2379–2407, (2011).

[8] T. Flåtten, A. Morin and S. T. Munkejord, Wave propagation in multicomponent flow models, *SIAM J. Appl. Math.* **70**, 2861–2882, (2010).

[9] A. Goldshtein and M. Shapiro, Mechanics of collisional motion of granular materials: Part 1. General hydrodynamic equations, *J. Fluid Mech.* **282**, 75–114, (1995).

[10] P.K. Haff, Grain flow as a fluid mechanical phenomenon, *J. Fluid Mech.* **134**, 401–430, (1983).

[11] J. Hersch, Contribution a la méthode des équations aux différences, *Z. Angew. Math. Phys.* **9**, 129–180, (1958).

[12] M. Hochbruck, C. Lubich and H. Selhofer, Exponential integrators for large systems of differential equations, *SIAM J. Sci. Comput.* **19**, 1552–1574, (1998).

[13] M. Hochbruck and A. Ostermann, Explicit exponential Runge–Kutta methods for semilinear parabolic problems, *SIAM J. Numer. Anal.* **43**, 1069–1090, (2005).

[14] M. Hochbruck and A. Ostermann, Exponential Runge–Kutta methods for parabolic problems, *Appl. Numer. Math.* **53**, 323–339, (2005).

[15] M. Hochbruck and A. Ostermann, Explicit integrators of Rosenbrock-type, *Oberwolfach Reports* **3**, 1107–1110, (2006).

[16] M. Hochbruck, A. Ostermann and J. Schweitzer, Exponential Rosenbrock-type methods, *SIAM. J. Numer. Anal.* **47**, 786–803, (2009).

[17] M. Hochbruck and A. Ostermann, Exponential integrators, *Acta Numerical* **19**, 209–286, (2010).

[18] H. Holden, K. H. Karlsen, K.-A. Lie and N. H. Risebro, *Splitting methods for partial differential equations with rough solutions*, EMS Series of Lectures in Mathematics, EMS Publishing House, Zürich, (2010).

[19] S. Jin, Runge–Kutta methods for hyperbolic conservation laws with stiff relaxation terms, *J. Comput. Phys.* **122**, 51–67, (1995).

[20] H. Kamath and X. Du, A roe-average algorithm for a granular-gas model with non-conservative terms, *J. Comput. Phys.* **228**, 8187–8202, (2009).

[21] D. Kincaid and W. Cheney, *Numerical analysis: mathematics of scientific computing.* American Mathematical Society, Providence, RI, USA, (2009).

[22] S. Krogstad, Generalized integrating factor methods for stiff PDEs, *J. Comput. Phys.* **203**, 72–88, (2005).
[23] J. D. Lawson, Generalized Runge–Kutta processes for stable systems with large Lipschitz constants, *SIAM J. Numer. Anal.* **4**, 372–380, (1967).
[24] T.-P. Liu, Hyperbolic conservation laws with relaxation, *Commun. Math. Phys.* **108**, 153–175, (1987).
[25] S. T. Munkejord, A numerical study of two-fluid models with pressure and velocity relaxation, *Adv. Appl. Math. Mech.* **2**, 131–159, (2010).
[26] R. Natalini, Recent results on hyperbolic relaxation problems. Analysis of systems of conservation laws, in *Chapman & Hall/CRC Monogr. Surv. Pure Appl. Math.*, 99, Chapman & Hall/CRC, Boca Raton, FL, 128–198, (1999).
[27] S. P. Nørsett, An A-stable modification of the Adams-Bashforth methods, in *Conference on the Numerical Solution of Differential Equations. Lecture Motes in Mathematics*, 109, Springer, 214–219, (1969).
[28] S. Osher, Convergence of generalized MUSCL schemes, *SIAM J. Numer. Anal.* **22**, 947–961, (1985).
[29] L. Pareschi and G. Russo, Implicit-explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation, *J. Sci. Comput.* **25**, 129–155, (2005).
[30] M. Pelanti, F. Bouchut and A. Mangeney, A Roe-type scheme for two-phase shallow granular flows over variable topography, *ESAIM: M2AN* **42**, 851–885, (2008).
[31] E. C. Rericha, C. Bizon, M. D. Shattuck and H. L. Swinney, Shocks in supersonic sand, *Phys. Rev. Lett.* **88**, 14302, (2001).
[32] R. Saurel, F. Petitpas and R. Abgrall, Modelling phase transition in metastable liquids: application to cavitating and flashing flows, *J. Fluid Mech.* **607**, 313–350, (2008).
[33] S. Serna and A. Marquina, Capturing shock waves in inelastic granular gases, *J. Comput. Phys.* **209**, 787–795, (2005).
[34] G. Strang, On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.* **5**, 506–517, (1968).
[35] E. F. Toro, MUSTA: A multi-stage numerical flux, *Appl. Numer. Math.* **56**, 1464–1479, (2006).
[36] B. van Leer, Towards the ultimate conservative difference scheme V. A second-order sequel to Godunov's method, *J. Comput. Phys.* **32**, 101–136, (1979).
[37] A. Zein, M. Hantke and G. Warnecke, Modeling phase transition for compressible two-phase flows applied to metastable liquids, *J. Comput. Phys.* **229**, 2964–2998, (2010).