ASYMPTOTIC ANALYSIS OF A STRUCTURE-PRESERVING INTEGRATOR FOR DAMPED HAMILTONIAN SYSTEMS

Adrian Viorel*

Babeş-Bolyai University Cluj-Napoca
Str. Kogalniceanu 3, 400084 Cluj-Napoca, Romania

Cristian D. Alecsa

Romanian Institute of Science and Technology
400022 Cluj-Napoca, Romania and
Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy
400320 Cluj-Napoca, Romania

Titus O. Pinţa

Institute for Numerical and Applied Mathematics, University of Göttingen
Lotzestr. 16-18, 37083 Göttingen, Germany

(Communicated by J. M. Sanz-Serna)

Abstract. The present work deals with the numerical long-time integration of damped Hamiltonian systems. The method that we analyze combines a specific Strang splitting, that separates linear dissipative effects from conservative ones, with an energy-preserving averaged vector field (AVF) integrator for the Hamiltonian subproblem. This construction faithfully reproduces the energy-dissipation structure of the continuous model, its equilibrium points and its natural Lyapunov function. As a consequence of these structural similarities, both the convergence to equilibrium and, more interestingly, the energy decay rate of the continuous dynamical system are recovered at a discrete level. The possibility of replacing the implicit AVF integrator by an explicit Störmer-Verlet one is also discussed, while numerical experiments illustrate and support the theoretical findings.

1. **Introduction.** Optimization has played an important role in the development of modern science ever since Fermat and Maupertuis have laid its foundations. Nowadays, this fundamental paradigm plays a central role in some of the most challenging fields of applied mathematics such as machine learning or image processing, with the difference that emphasis lays on an algorithmic perspective.

In these fields, second order dissipative dynamical systems have become quite fashionable during the last couple of decades, as part of a revival of Polyak's [34] seminal idea that optimization can be connected to dissipative systems by considering a damped mechanical system whose potential energy is precisely the function to be minimized and taking advantage of the system's natural tendency to converge to equilibrium, i. e., to minimize the potential energy. The most telling example of

such a mechanical system is that of a heavy ball sliding, under the action of gravity, along a profile defined by the function to be minimized. This is the celebrated heavy ball with friction dynamical system [2].

Independently of their relevance to optimization, dissipative dynamical systems governed by second order evolution equations have been studied in connection to various applications (see for example [18], [22] or [42]), especially related to damped wave or beam equations. In the more challenging, nonconvex case, the asymptotic behavior of such dissipative systems has been analyzed by Haraux and Jendoubi [20], [21], who pioneered the use of the Łojasiewicz inequality as a main ingredient for obtaining energy decay rates in the long time asymptotic regime. Their original results have been complemented not long ago by Bégout, Bolte and Jendoubi [4] who were able to derive, in a general case, convergence rates for trajectories based on a quasi-gradient system approach (cf. also [23]).

In this context, the present contribution is concerned with the asymptotic analysis of a structure-preserving scheme based on reversible-irreversible splitting, which faithfully reproduces the long-time decay to equilibrium of the continuous system.

The asymptotic behavior of splitting methods has been investigated recently by Dujardin and Lafitte [10]. However, we deviate from their approach in the sense that we do not consider the asymptotic error, i.e., the difference between exact and numerical asymptotic states, as defined in [10], rather, we restrict our attention to an integrator whose asymptotic states inherently coincide with those of the continuous system. This, together with correctly reproducing the energy-structure of the problem and its asymptotic energy decay rate are the main objectives of this work and they relate to one of the fundamental questions of geometric numerical integration: *"How can numerical methods be constructed that respect the geometry of the problem at hand?"* (see [15]).

When choosing the numerical scheme, we take a (Strang) splitting approach, which allows semiflows (subproblems) to be chosen in a natural way such that linear dissipative effects are completely uncoupled from nonlinear conservative ones. The ensuing discrete dynamical system preserves both the Lyapunov function and the equilibrium points of its continuous counterpart. This approach is not new as splitting methods have been used to exactly preserve the conformal symplectic nature of the system (see [30, 31]). In most cases, the nonlinear conservative subproblem, a Hamiltonian evolution system, cannot be solved explicitly, but a variety of effective geometric numerical integrators are available. In [5] both implicit midpoint methods and Störmer-Verlet methods are shown to lead to second order full discretizations that preserve dissipation of symplecticity. Under sufficient smallness assumptions on the step size as well as the damping coefficient, previous work of Modin and Söderlind [32] had used backward error analysis to show that energy dissipation is correct up to an exponentially small error, while relative equilibria are nearly preserved. Very recently, the same approximation ideas have been investigated in the more general setting of dissipative systems for which an underlying thermodynamic (GENERIC) structure exists by Shang and Öttinger [38]. Their findings highlight, along with accuracy, the excellent structural properties of the approach.

In this work, among the existing Hamiltonian integrators, we choose the energy-conserving averaged vector field (AVF) method of Quispel and McLaren [36]. This may not seems to be the most natural choice due to the implicit character of AVF, but exact energy conservation plays a central role in our asymptotic analysis. This

choice has also been taken into account in [38] where Verlet splitting integrators are juxtaposed to a classical Runge-Kutta method and also to an AVF-based method.

From an optimization perspective, the AVF idea has been analyzed recently by Ehrhardt et. al. [11] (see also [17]) as one of the possible discrete gradient geometric integrators that approximate a continuous gradient flow inheriting its dissipation property and convergence to equilibrium. There is also a more general increased interest of the optimization community in geometric integrators and the connections between their field and the numerical analysis of evolution equations, as shown by the large number of diverse recent contributions like [9, 14, 24, 29, 39, 43]. We must point out however, that the present analysis concentrates on nonconvex energies while all the above references are bound to the convex case and not directly comparable. For recent contributions dealing with both continuous and discrete dynamics in the nonconvex stetting, we refer to [6], [26] or [1].

The paper is organized as follows. In Section 2, we provide a detailed analysis of the continuous problem. While not new, these results have been included since they provide a blueprint for the constructions in Sections 3 and 4 where the properties of the continuous system are replicated at a discrete level. As a first step towards a full discretization, in Section 3, a conservative-dissipative Strang splitting that correctly reproduces the energy-dissipation structure of the problem is proposed, while in Section 4 the Hamiltonian subproblem is discretized giving rise to a composite geometric integrator. The analysis of this AFV splitting integrator, still carried out in Section 4, proves that for the generated discrete dynamical system convergence to equilibrium takes place with the same energy decay rate as for its continuous counterpart. An explicit alternative to the AVF integrator is discussed in Section 5. Finally, numerical experiments supporting all these results are presented in Section 6.

2. **Analysis of the continuous model.** We consider the second-order gradient system $\ddot{u} + 2\dot{u} = -\nabla E(u)$ or, equivalently,

$$\begin{cases} \dot{u} & = & v, \\ \dot{v} & = & -2v - \nabla E(u). \end{cases} \tag{1}$$

which can be understood as describing the damped oscillations of a particle under the action of a conservative force derived form the potential E. The choice of the damping coefficient is obviously arbitrary but will help simplifying later computations without changing the nature of the results. Owing to its mechanical structure, arguably the most important feature of the system is its energy balance equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|v(t)\|^2 + E(u(t))\right) = -2\|v(t)\|^2, \tag{2}$$

according to which the total (kinetic plus potential) energy (Hamiltonian) of the system,

$$H(u(t), v(t)) = \frac{1}{2}\|v(t)\|^2 + E(u(t)) \tag{3}$$

is decreasing along solutions. Frictional forces will diminish the energy of the system, and hence the amplitude of its oscillations, until eventually the system approaches a state of mechanical equilibrium (u^*, v^*) at which both frictional and potential forces vanish, that is,

$$v^* = 0 \quad \text{and} \quad \nabla E(u^*) = 0.$$

Given that the potential energy E has C^2 regularity, a standard stability analysis for the above equilibria shows that pairs $(u^*_{max}, 0)$ where u^*_{max} is a maximum of E are unstable, while pairs $(u^*_{min}, 0)$ are exponentially stable when u^*_{min} is a minimizer of the potential energy. Thus, from an optimization perspective, the dynamical system (1) and its discrete counterparts are expected to provide a viable approach for finding local minima of an objective function.

Dissipation of energy assures the existence of global in time solutions to (1) under very natural assumptions concerning the regularity and coercivity, i.e., $E(u) \to \infty$ as $\|u\| \to \infty$, of the potential E. However, the energy balance equation alone does not provide sufficient information in order to fully characterize the rate of convergence towards equilibrium, such that additional information concerning the shape of the potential E at critical points u^* is necessary. The property that enables the derivation of convergence rates is due, originally, to Łojasiewicz [27].

Definition 2.1. Let $E : \mathbb{R}^N \to \mathbb{R}$ be a differentiable function. We say that E has the Łojasiewicz property at a critical point $u^* \in \mathrm{crit}E$ if there exists an open neighborhood V_{u^*} of u^* and constants $\theta \in (0,1)$, $C > 0$ such that

$$|E(u) - E(u^*)|^\theta \leq C\|\nabla E(u)\| \quad \text{for all} \quad u \in V_{u^*}. \tag{4}$$

Based on fundamental results of Haraux and Jendoubi [20, 21] as well as Bégout, Bolte and Jendoubi [4], the analysis of the dissipative dynamical system (1) can be summarized in the following theorem which will serve as guideline for the analysis of the numerical scheme, that we carry out in Section 4.

Theorem 2.2. *Let $E : \mathbb{R}^N \to \mathbb{R}$ be twice continuously Fréchet differentiable and coercive. Then, for any initial condition $(u_0, v_0) \in \mathbb{R}^N \times \mathbb{R}^N$ there exists a unique, global in time solution $(u(t), v(t)) \in C^1([0, \infty), \mathbb{R}^N \times \mathbb{R}^N)$ of the Cauchy problem associated to (1) and all such trajectories converge to equilibrium in the sense that the limit $H^*(u_0, v_0) = \lim_{t\to\infty} H(u(t), v(t))$ exists and*

$$\mathrm{dist}\left((u(t), v(t)), \mathcal{E}\right) \to 0 \quad as \quad t \to \infty,$$

where $\mathcal{E} = \left\{(u^*, 0) \in \mathbb{R}^N \times \mathbb{R}^N : \nabla E(u^*) = 0\right\}$ is the set of equilibria of the system. Furthermore, if $\mathcal{E}$ consists only of isolated equilibria, then

$$u(t) \to u^* \quad as \quad t \to \infty$$

for some $(u^*, 0) \in \mathcal{E}$ and if additionally the system's potential energy E has the Łojasiewicz property with $\theta = 1/2$ at u^* then there exists $\varepsilon > 0$ small enough such that E decays exponentially starting with a large enough time $t_0 > 0$

$$E(u(t)) - E(u^*) \leq Ce^{-\varepsilon(t-t_0)}. \tag{5}$$

Exponential decay is achieved also in norm $\|u(t) - u^*\| \leq Ce^{-\omega(t-t_0)}$.

Proof. We divide the proof into several steps.

Step 1. Local (in time) existence of solutions follows from a contraction principle argument based on the fact that (1) has a locally Lipschitz continuous right hand side (due to the regularity of E).

Step 2. Local solutions $(u(t), v(t)) \in C^1([0, t_{max}), \mathbb{R}^N \times \mathbb{R}^N)$ can be extended arbitrarily in time based on the coercivity of E. In fact, along solutions we have that the total energy is decreasing rendering finite time blow up impossible (see [37] for the general discussion).

Step 3. Convergence to equilibrium is a consequence of LaSalle's invariance principle (see Theorems 9.2.3, 9.2.7 and Corollary 9.2.9 in [7] which apply directly) as H is a strict Lyapunov function.

Qwing to the convergence to equilibrium, the decrease property of H and on the fact that $E(u(t)) \leq H(u(t), v(t))$, there exists a large enough time $t_0 > 0$ after which the trajectory remains in a small enough energy neighborhood of the equilibrium point and, furthermore, $u(t)$ is confined to a sublevel set of E that is itself contained in the neighborhood V_{u^*} in which the Łojasiewicz property holds. To derive the decay estimate (5), we proceed in two steps.

Step 4. First, we prove that, for $t \geq t_0$ large enough such that the Łojasiewicz inequality holds and $\varepsilon > 0$ small enough, the modified energy

$$\widetilde{H}(u, v) = \frac{1}{2} \|v\|^2 + (E(u) - E(u^*)) + \varepsilon \langle \nabla E(u),\ v \rangle_{\mathbb{R}^N} \qquad (6)$$

satisfies a decay inequality

$$\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}t}(u, v) \leq -\varepsilon \widetilde{H}(u, v) = \frac{\varepsilon}{2} \|v\|^2 + \varepsilon (E(u) - E(u^*)) + \varepsilon^2 \langle \nabla E(u),\ v \rangle_{\mathbb{R}^N}, \qquad (7)$$

in which the explicit time-dependence of u, v has been omitted for brevity reasons. This inequality later yields exponential energy decay along solutions. Indeed, using energy dissipation equation and and the second equation in (1) we have

$$\begin{aligned}
\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}t}(u, v) &= -2 \|v\|^2 + \varepsilon \langle \nabla^2 E(u)\dot{u},\ v \rangle_{\mathbb{R}^N} + \varepsilon \langle \nabla E(u),\ \dot{v} \rangle_{\mathbb{R}^N} \\
&= -2 \|v\|^2 + \varepsilon \langle \nabla^2 E(u)v,\ v \rangle_{\mathbb{R}^N} - \varepsilon \langle \nabla E(u),\ 2v \rangle_{\mathbb{R}^N} - \varepsilon \|\nabla E(u)\|^2.
\end{aligned}$$

As the trajectory remains contained in a closed bounded set

$$\varepsilon \langle \nabla^2 E(u)v,\ v \rangle_{\mathbb{R}^N} \leq \varepsilon M \|v\|^2$$

such that after forcing the desired ε^2 and applying Young's inequality to the reminder term

$$(\varepsilon^2 - 2\varepsilon) \langle \nabla E(u),\ v \rangle_{\mathbb{R}^N} = \varepsilon \langle \nabla E(u),\ (\varepsilon - 2)v \rangle_{\mathbb{R}^N} \leq \frac{\varepsilon}{2} \|\nabla E(u)\|^2 + \frac{\varepsilon(\varepsilon - 2)^2}{2} \|v\|^2$$

we arrive at

$$\begin{aligned}
\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}t}(u, v) &\leq -2 \|v\|^2 + \varepsilon M \|v\|^2 - \varepsilon^2 \langle \nabla E(u),\ v \rangle_{\mathbb{R}^N} \\
&\quad + \frac{\varepsilon}{2} \|\nabla E(u)\|^2 + \frac{\varepsilon}{2}(2 - \varepsilon)^2 \|v\|^2 - \varepsilon \|\nabla E(u)\|^2.
\end{aligned}$$

After combining the $\|\nabla E(u)\|^2$ terms and using the Łojasiewicz inequality (4), one finally obtains

$$\frac{\mathrm{d}\widetilde{H}}{\mathrm{d}t}(u, v) \leq -(2 - \varepsilon M - \frac{\varepsilon}{2}(2 - \varepsilon)^2) \|v\|^2 - \varepsilon (E(u) - E(u^*)) - \varepsilon^2 \langle \nabla E(u),\ v \rangle_{\mathbb{R}^N},$$

which gives exactly (7) if ε is taken such that $\frac{\varepsilon}{2} \leq 2 - \varepsilon M - \frac{\varepsilon}{2}(2 - \varepsilon)^2$.

Step 5. Obviously, the differential inequality (7) implies the exponential decay of $\widetilde{H}$, i.e.,

$$\widetilde{H}(u(t), v(t)) \leq e^{-\varepsilon(t - t_0)} \widetilde{H}(u(t_0), v(t_0)) \quad \text{for all} \quad t \geq t_0.$$

However, since $\langle \nabla E(u(t)),\ v(t)\rangle_{\mathbb{R}^N} = \frac{\mathrm{d}}{\mathrm{d}t}(E(u(t)) - E(u^*))$ the above decay inequality rewrites as

$$\frac{1}{2}\|v\|^2 + (E(u) - E(u^*)) + \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}(E(u(t)) - E(u^*)) \leq e^{-\varepsilon(t-t_0)}\widetilde{H}(u(t_0), v(t_0))$$

and this provides a differential inequality for $E(u(t)) - E(u^*)$

$$\frac{\mathrm{d}}{\mathrm{d}t}(E(u(t)) - E(u^*)) \leq -\frac{1}{\varepsilon}(E(u(t)) - E(u^*)) + e^{-\varepsilon(t-t_0)}\widetilde{H}(u(t_0), v(t_0)).$$

In view of this inequality, which can be integrated by the variation of constants formula (and with the notation $\widetilde{H}_0 = \widetilde{H}(u(t_0), v(t_0))$)

$$E(u(t)) - E(u^*) = e^{-\frac{1}{\varepsilon}(t-t_0)}(E(u(t_0)) - E(u^*)) + \frac{\widetilde{H}_0}{\varepsilon}\int_{t_0}^{t} e^{-\frac{1}{\varepsilon}(t-s)}e^{-\varepsilon(s-t_0)}\,\mathrm{d}s\,.$$

After explicitly computing the integral on the right hand side the desired decay estimate is obtained under a $\varepsilon \ll 1$ smallness assumption

$$E(u(t)) - E(u^*) \leq e^{-\frac{1}{\varepsilon}(t-t_0)}z(t_0) + \frac{\widetilde{H}_0}{1 - \varepsilon^2}\left(e^{-\varepsilon(t-t_0)} - e^{-\frac{1}{\varepsilon}(t-t_0)}\right).$$

Step 6. For the proof of exponential norm convergence we refer to [4].   □

Remark 1. The eventual exponential decay of the potential energy (5) actually implies the eventual exponential decay of the total energy H.

3. **The Strang splitting semidiscretization.** Originally developed in a linear setting, operator splitting methods are based on the idea of replacing the more complicated (linear) problem $\dot{z} = (A + B)z$ by two simpler (sub)problems $\dot{z} = Az$ and $\dot{z} = Bz$ which are solved iteratively. In the late 19th century, S. Lie proposed replacing the exact solution operator $e^{h(A+B)}$ of the problem by the approximation $e^{hA}e^{hB}$, the so called Lie splitting. In the present work, we will rather focus on the second order, symmetrized version of the Lie scheme introduced by G. Strang [40] (see also [28]) $e^{h(A+B)} \approx e^{\frac{h}{2}A}e^{hB}e^{\frac{h}{2}A}$. In a very general setting, Hansen, Kramer and Ostermann [19] have shown that the linearity of the equation is not essential, as the semilinear abstract Banach space evolution equation

$$\dot{z} = Az + f(z) \tag{8}$$

can be treated by Strang splitting in exactly the same manner as the linear problem, replacing only e^{hB} with the solution operator, i.e., nonlinear semigroup, $U(h)$ of the nonlinear subproblem $\dot{z} = f(z)$.

The system (1) is semilinear and fits nicely in the framework of Hansen, Kramer and Ostermann. Moreover the splitting method can be applied in such a way that the conservative and dissipative parts of the system are separated. We replace the original problem (1) by two subproblems

$$\begin{cases} \dot{u} &= 0, \\ \dot{v} &= -2v, \end{cases} \tag{9}$$

and

$$\begin{cases} \dot{u} &= v, \\ \dot{v} &= -\nabla E(u), \end{cases} \tag{10}$$

one being linear, dissipative and explicitly solvable while the second is nonlinear and conservative.

This is not the only possible choice for the two subproblems. It is, actually, a special case ($\beta = 1$) of the more general parameterized family of decompositions featuring in the context damped of semilinear wave equation in [35], namely,

$$\begin{cases} \dot{u} & = & (1-\beta)v, \\ \dot{v} & = & -2v, \end{cases} \quad \text{and} \quad \begin{cases} \dot{u} & = & \beta v, \\ \dot{v} & = & -\nabla E\,(u)\,, \end{cases}$$

with $\beta \in [0,1]$. However, the choice $\beta = 1$, that we will concentrate on, seems to be most appropriate for separating conservative from dissipative effects.

Adopting matrix notations and using $\{U(t)\}_{t\in\mathbb{R}}$ as a notation for the nonlinear (semi)group generated by the Hamiltonian subproblem (10), we construct the discrete dynamical system

$$\left[\begin{array}{c} u_{n+1} \\ v_{n+1} \end{array} \right] = S_h \left[\begin{array}{c} u_n \\ v_n \end{array} \right], \qquad S_h = \left[\begin{array}{cc} \mathrm{I} & 0 \\ 0 & e^{-h}\mathrm{I} \end{array} \right] U(h) \left[\begin{array}{cc} \mathrm{I} & 0 \\ 0 & e^{-h}\mathrm{I} \end{array} \right]. \qquad (11)$$

Before embarking on a detailed discussion of this approximation's properties let us recall the discrete version of LaSalle's invariance principle that plays a central role in the subsequent asymptotic analysis. We consider a continuous map $T : \mathbb{R}^N \to \mathbb{R}^N$ and the associated discrete dynamical system

$$x_{n+1} = T(x_n), \qquad (12)$$

while $\mathcal{E}_T = \left\{ x^* \in \mathbb{R}^N : x^* = T(x^*) \right\}$ denotes the set of equilibrium points of this system.

Definition 3.1. We say that $\Phi : \mathbb{R}^N \to \mathbb{R}$ is a strict Liapunov function for (12) if Φ is continuous on $\mathbb{R}^N$ and $\Phi(T(x)) < \Phi(x)$ for all $x \in \mathbb{R}^N \setminus \mathcal{E}_T$.

The fundamental result concerning convergence to equilibrium for discrete dynamical systems goes back to J. P. LaSalle. We present an adapted version, based on [25], that aims at highlighting the similarity to the continuous case used in Section 2.

Theorem 3.2 (**LaSalle's invariance principle**). *Let $\Phi : \mathbb{R}^N \to \mathbb{R}$ be a coercive strict Liapunov function for the discrete dynamical system (12). Then, for any $x_0 \in \mathbb{R}^N$ the generated trajectory $x_n = T^n(x_0)$ is bounded and*

(i) there exists $\Phi^* \in \mathbb{R}$ such that $\lim_{n\to\infty} \Phi(x_n) = \Phi^*$ while

$$\mathrm{dist}\,(x_n, \mathcal{E}_T) \to 0 \quad as \quad n \to \infty.$$

(ii) Furthermore, if the set of equilibria is discrete, then for some $x^* \in \mathcal{E}_T$

$$x_n \to x^* \quad as \quad n \to \infty.$$

The following result shows that the Strang semidiscretization inherits the convergence to equilibrium of the continuous system.

Theorem 3.3. *Let E be a coercive C^2 potential and assume that the conservative subproblem (10) does not admit periodic solutions of period equal to the step size $h > 0$ then*

(i) the equilibrium points (fixed points) of the discrete dynamical system coincide with those of the continuous system;

(ii) the total energy H is a strict Lyapunov function for the discrete dynamical system such that

(iii) *for any initial state* $(u_0, v_0) \in \mathbb{R}^N \times \mathbb{R}^N$ *the trajectories generated by* (11) *converge to equilibrium in the sense that the limit* $H^*(u_0, v_0) = \lim_{t \to \infty} H(u_n, v_n)$ *exists and*

$$\text{dist}\,((u_n, v_n), \mathcal{E}) \to 0 \quad as \quad n \to \infty.$$

Furthermore, if $\mathcal{E}$ *consists only of discrete equilibria, then for some* $(u^*, 0) \in \mathcal{E}$

$$u_n \to u^* \quad as \quad n \to \infty.$$

Proof. If (u^*, v^*) is an equilibrium point for the discrete dynamical system generated by S_h and $\{U(t)\}_{t \in \mathbb{R}}$ is the nonlinear (semi)group generated by the Hamiltonian equation (10) then

$$U(h) \begin{bmatrix} u^* \\ e^{-h}v^* \end{bmatrix} = \begin{bmatrix} u^* \\ e^h v^* \end{bmatrix} \tag{13}$$

and, as the Hamiltonian flow conserves the total energy, $H(u^*, e^{-h}v^*) = H(u^*, e^h v^*)$. For $h > 0$, this can hold only if $v^* = 0$. Reverting to (13) bearing in mind the assumption that $U(h)$ has no periodic points, the only possibility left is that $(u^*, 0)$ is an equilibrium point for (10), that is, $\nabla E(u^*) = 0$.

To prove that H is a strict Lyapunov function, we analyze the change of the total energy along each intermediate stage of the splitting. The advantage of this splitting choice becomes now obvious as the dissipative splitting step only affects (decreases) the kinetic energy while the conservative splitting step conserves the total energy. Denoting

$$\begin{bmatrix} u_{n+1/3} \\ v_{n+1/3} \end{bmatrix} = \begin{bmatrix} \mathrm{I} & 0 \\ 0 & e^{-h}\mathrm{I} \end{bmatrix} \begin{bmatrix} u_n \\ v_n \end{bmatrix}, \quad \begin{bmatrix} u_{n+2/3} \\ v_{n+2/3} \end{bmatrix} = U(h) \begin{bmatrix} u_{n+1/3} \\ v_{n+1/3} \end{bmatrix}$$

we have

$$H(u_{n+1}, v_{n+1}) \le H(u_{n+\frac{2}{3}}, v_{n+\frac{2}{3}}) = H(u_{n+\frac{1}{3}}, v_{n+\frac{1}{3}}) \le H(u_n, v_n),$$

with equality holding only if $v_{n+\frac{k}{3}} = 0, k = 0, \ldots 3$, that is, only for exact equilibria $(u^*, 0) \in \mathcal{E}$. The asymptotic behavior of the dynamical systems now follows from the discrete version of LaSalle's invariance principle presented in Theorem 3.2. □

Remark 2 (accuracy). A straightforward application of results in [19] assures that the scheme is second-order, for sufficiently regular E (i.e. C^3 regularity).

4. **Analysis of the averaged vector field splitting numerical integrator.** Taking the Strang semidiscretization as a starting point, we observe that of the ensuing subproblems, the nonlinear Hamiltonian system is by far the more challenging one, as the dissipative part is exactly solvable. In view of its structure, we use the energy-preserving numerical integrator introduced by Quispel and McLaren [36] for discretizing (10), that is,

$$\begin{cases} u_{n+1} & = & u_n + \frac{h}{2}(v_{n+1} + v_n) \\[2mm] v_{n+1} & = & v_n - h \int_0^1 \nabla E((1-\alpha)u_n + \alpha u_{n+1})\, \mathrm{d}\alpha\,. \end{cases} \tag{14}$$

This Averaged Vector Field (AVF) method is known to be a second-order discretization of Hamiltonian systems (cf. [36]). Further, it is implicit and requires anti-derivatives. For most of the common Hamiltonians, such as polynomial ones, the anti-derivatives can be computed explicitly, or this drawback can be circumvent by replacing the AVF method with a (possibly higher-order) energy-preserving collocation method (see Hairer [16]). Notably, in the case of quadratic Hamiltonians, i.

e., linear Hamiltonian evolution equations, the AVF method is just the well-known trapezoidal discretization, see [8].

Denoting the discrete (nonlinear) AVF evolution operator generated by (14) with U_h^{AVF}, the full Strang splitting AVF discretization of the dissipative system (1) is given by the evolution operator

$$S_h^{AVF} = \begin{bmatrix} I & 0 \\ 0 & e^{-h}I \end{bmatrix} U_h^{AVF} \begin{bmatrix} I & 0 \\ 0 & e^{-h}I \end{bmatrix}$$

and reads

$$\begin{cases} u_{n+1} &= u_n + \frac{h}{2}(e^h v_{n+1} + e^{-h} v_n) \\ e^h v_{n+1} &= e^{-h} v_n - h \int_0^1 \nabla E((1-\alpha)u_n + \alpha u_{n+1}) \, d\alpha. \end{cases} \tag{15}$$

The main result of the present work is the following

Theorem 4.1. *Let us consider the discrete dynamical system generated by the numerical scheme* (15), *that is, the dynamical system generated by the operator* S_h^{AVF}. *If the potential E is of class C^2 and coercive then the following statements hold for any step size $h > 0$:*

(i) *the equilibrium points of the discrete dynamical system coincide with those of the continuous system;*

(ii) *the total energy H is a strict Lyapunov function for the discrete dynamical system and*

(iii) *for any initial data (u_0, v_0) the asymptotic behavior of the system is governed by*

$$\text{dist}\left([u_n, v_n]^T, \mathcal{E}\right) \to 0 \quad as \quad n \to \infty$$

where $[u_n, v_n]^T = (S_h^{AVF})^n [u_0, v_0]^T$. Moreover, if $\mathcal{E}$ consists only of isolated equilibria (or equivalently, if all critical points of E are isolated), then

$$u_n \to u^* \quad as \quad n \to \infty$$

for some u^* with $\nabla E(u^*) = 0$.

(iv) *Further, if the Lojasiewicz condition holds at u^*, then there exist $\varepsilon > 0$ and $n_0 > 0$ such that*

$$E(u_n) - E(u^*) \leq C \left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}} \right)^{n - n_0}, \quad for \ n \geq n_0. \tag{16}$$

Proof. The fact that the numerical scheme (15) has the same equilibria as the continuous system is straightforward to check.

Proving that H is a Lyapunov function for the discrete dynamical system generated by the numerical scheme relies on the same arguments as in the proof of Theorem 3.3 sice the AVF step conserves the total energy. Convergence to equilibrium follows from LaSalle's discrete invariance principle.

In order to obtain a rate of decay for the potential energy, we start by deducing a discrete version of the energy balance equation. Let us first observe that, bearing to the AVF structure (cf. [36]),

$$\left\langle \int_0^1 \nabla E\left((1-\alpha)u_n + \alpha u_{n+1}\right) d\alpha, \ u_{n+1} - u_n \right\rangle_{\mathbb{R}^N} = E(u_{n+1}) - E(u_n). \tag{17}$$

Rewriting the second equation in (15) as

$$e^h v_{n+1} - e^{-h} v_n = -h \int_0^1 \nabla E \left((1-\alpha) u_n + \alpha u_{n+1} \right) d\alpha$$

and taking the inner product with $\frac{1}{h}(u_{n+1} - u_n) = \frac{1}{2}(e^h v_{n+1} + e^{-h} v_n)$, this being the first equation in (15), yields the discrete energy balance

$$\frac{e^{2h}}{2} \|v_{n+1}\|^2 + E(u_{n+1}) = \frac{e^{-2h}}{2} \|v_n\|^2 + E(u_n). \tag{18}$$

Here the dissipation, a loss of kinetic energy, is hidden in the weights $e^{-2h} \leq 1 \leq e^{2h}$ that multiply the kinetic energy terms. To derive an estimate of the decrease in total energy we use

$$e^{2h} \geq 1 + 2h + 2h^2 \quad \text{and} \quad e^{-2h} \leq 1 - 2h + 2h^2 \tag{19}$$

such that (18) becomes

$$(1+2h^2)\left(\frac{1}{2}\|v_{n+1}\|^2 - \frac{1}{2}\|v_n\|^2 \right) + E(u_{n+1}) - E(u_n) \leq -h(\|v_{n+1}\|^2 + \|v_n\|^2), \tag{20}$$

meaning that the technically useful adjusted total energy

$$H^h(u_n, v_n) = (1+2h^2)\frac{1}{2}\|v_n\|^2 + E(u_n) \tag{21}$$

is also a strict Lyapunov function. The estimate (20) is sufficient for deducing the desired behavior as it is a discrete counterpart of (2), which is recovered in the $h \to 0$ limit.

For the sake of simplicity, the analysis is carried out only in the case $E(u^*) = 0$. Our strategy is to establish discrete counterparts for Steps 4 and 5 in the proof of Theorem 2.2. Accordingly, we introduce the discrete version of the modified energy $\widetilde{H}$ in (6)

$$\widetilde{H}_n^{n+1} = \frac{H^h(u_{n+1}, v_{n+1}) + H^h(u_n, v_n)}{2} + \varepsilon \frac{E(u_{n+1}) - E(u_n)}{h}. \tag{22}$$

Using the AVF property (17) and convenient shorthand notations for the discrete gradient and energies

$$\overline{\nabla} E_n^{n+1} = \int_0^1 \nabla E \left((1-\alpha) u_n + \alpha u_{n+1} \right) d\alpha, \tag{23}$$

$$H_n^h = H^h(u_n, v_n) \quad \text{and} \quad E_n = E(u_n) \tag{24}$$

we can write $\widetilde{H}_n^{n+1}$ in a form that highlights the similarity to (6) namely

$$\widetilde{H}_n^{n+1} = \frac{H_{n+1}^h + H_n^h}{2} + \varepsilon \left\langle \overline{\nabla} E_n^{n+1}, \frac{1}{2}(e^h v_{n+1} + c e^{-h} v_n) \right\rangle_{\mathbb{R}^N}. \tag{25}$$

When comparing to the continuous case one can see that the total energy at time t has been replaced by the average modified energy of two consecutive steps. Also, the gradient ∇E of the potential energy has been replaced by its discrete AVF version, while as a substitute of the velocity $v(t)$ we have a weighted average of consecutive velocities.

A crucial, but rather technical, part of the proof that we postpone to Appendix A consists in deriving that for a large enough $n_0 \in \mathbb{N}$

$$\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \leq -\frac{\varepsilon h}{2} \left(\widetilde{H}_n^{n+1} + \widetilde{H}_{n-1}^n \right), \quad \text{for all} \quad n > n_0. \tag{26}$$

This discrete version of (7) assures the decay of $\widetilde{H}_n^{n+1}$

$$\widetilde{H}_n^{n+1} \leq \left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}}\right)^{n-n_0} \widetilde{H}_{n_0}^{n_0+1}, \quad \text{for all} \quad n > n_0. \tag{27}$$

Returning to (22), one can see that the decay of $\widetilde{H}_n^{n+1}$ translates, after discarding the positive kinetic energy terms, into a difference inequality for the values of the potential E

$$\frac{E_{n+1} + E_n}{2} + \varepsilon \frac{E_{n+1} - E_n}{h} \leq \left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}}\right)^{n-n_0} \widetilde{H}_{n_0}^{n_0+1}.$$

This is an inhomogeneous difference inequality

$$E_{n+1} \leq \left(\frac{1 - \frac{h}{2\varepsilon}}{1 + \frac{h}{2\varepsilon}}\right) E_n + \frac{h}{\varepsilon} \frac{\widetilde{H}_{n_0}^{n_0+1}}{1 + \frac{h}{2\varepsilon}} \left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}}\right)^{n-n_0}.$$

that can be solved (cf. [13] for a detailed discussion of discrete Gronwall-type inequalities).Straightforward computations lead to

$$E_n \leq \left(\frac{1 - \frac{h}{2\varepsilon}}{1 + \frac{h}{2\varepsilon}}\right)^{n-n_0} E_{n_0} + \frac{h}{\varepsilon} \frac{\widetilde{H}_{n_0}^{n_0+1}}{1 + \frac{h}{2\varepsilon}} \sum_{j=n_0}^{n-1} \left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}}\right)^{n-j-1} \left(\frac{1 - \frac{h}{2\varepsilon}}{1 + \frac{h}{2\varepsilon}}\right)^{j-n_0}.$$

Using the sum of a geometric progression we finally arrive at

$$E_n \leq \left(\frac{1 - \frac{h}{2\varepsilon}}{1 + \frac{h}{2\varepsilon}}\right)^{n-n_0} E_{n_0} + \frac{2 - \varepsilon h}{2(1 - \varepsilon^2)} \left(\left(\frac{1 - \frac{\varepsilon h}{2}}{1 + \frac{\varepsilon h}{2}}\right)^{n-n_0} - \left(\frac{1 - \frac{h}{2\varepsilon}}{1 + \frac{h}{2\varepsilon}}\right)^{n-n_0}\right) \widetilde{H}_{n_0}^{n_0+1}.$$

For $\varepsilon \ll 1$, the leading order term is exactly that in (16). $\qquad \square$

Remark 3 (accuracy). Combining results concerning the accuracy of the semidiscretization [19] with those for the AVF method [16] assures that the method (15) is second order, provided that E is three times continuously differentiable.

5. **Discussion of an explicit alternative to the AVF splitting.** In this section, we discuss the possibility of replacing (15) by an explicit scheme. Despite its convincing analytic features presented in the previous section, the AVF splitting discretization is implicit and one would rather rely on an explicit scheme. As a possible alternative, we consider the same splitting approach as before but with the energy-preserving averaged vector field solver replaced by a Störmer-Verlet symplectic integrator, that is, precisely the conformal symplectic scheme proposed by Bhatt, Floyd and Moore in [5]

$$\begin{cases} u_{n+1} &= u_n + h e^{-h} v_n - \frac{h^2}{2} \nabla E(u_n), \\ e^h v_{n+1} &= e^{-h} v_n - \frac{h}{2} \left(\nabla E(u_n) + \nabla E(u_{n+1})\right). \end{cases} \tag{28}$$

The equilibria of this discrete dynamical system coincide with those of the original continuous system.

As we will see below, the main drawback of the explicit approach is, unsurprisingly, related to its step size restriction as well as its failure to reproduce the monotonic energy decrease of the continuous system. However, provided that the step size h is small enough, if the potential E is convex and has globally Lipschitz continuous gradient one can retrace the same qualitative convergence to equilibrium

as for the AVF discretization. Notably, the Lipschitz condition on the gradient of the potential is standard in convex optimization [33], but also highly restrictive as its imposes an at most quadratic growth of E at infinity

To prove our claim, we start by forcing (28) into a form similar to that of (15). In this sense, observe that using second equation in (28)

$$u_{n+1} - u_n = \frac{h}{2}(e^h v_{n+1} + e^{-h} v_n) + \frac{h^2}{4}\left(\nabla E(u_n) + \nabla E(u_{n+1})\right) - \frac{h^2}{2}\nabla E(u_n).$$

Now, also forcing an average discrete gradient in the second equation, (28) can be rewritten as

$$\begin{cases} u_{n+1} &= u_n + \frac{h}{2}(e^h v_{n+1} + e^{-h} v_n) + \frac{h^2}{4}\left(\nabla E(u_n) - \nabla E(u_{n+1})\right) \\ e^h v_{n+1} &= e^{-h} v_n - h\overline{\nabla} E_n^{n+1} + h\left(\overline{\nabla} E_n^{n+1} - \frac{\nabla E(u_{n+1}) + \nabla E(u_n)}{2}\right). \end{cases} \tag{29}$$

In hindsight of the AVF energy property (17), taking the inner product with $(u_{n+1} - u_n)/h$ yields

$$\frac{e^{2h}}{2}\|v_{n+1}\|^2 - \frac{e^{-2h}}{2}\|v_n\|^2 + E(u_{n+1}) - E(u_n) =$$

$$\frac{h^2}{8}\left\langle \nabla E(u_n) + \nabla E(u_{n+1}),\ \nabla E(u_n) - \nabla E(u_{n+1})\right\rangle_{\mathbb{R}^N} \tag{30}$$

$$+ \left\langle \overline{\nabla} E_n^{n+1} - \frac{\nabla E(u_n) + \nabla E(u_{n+1})}{2},\ u_{n+1} - u_n\right\rangle_{\mathbb{R}^N}.$$

Again, as in the proof of Theorem 4.1, the dissipated energy is hidden in the exponential weights of the kinetic energy terms and can be made explicit by (19). On the other hand, both error terms on the right hand side of the above energy balance equation can be masked (see Appendix B. for details) by dissipative effects provided that the step size conditions

$$hL < 1 \quad \text{and} \quad h < 4 \tag{31}$$

hold. The ensuing inequality

$$(1 + 2h^2)\left(\frac{1}{2}\|v_{n+1}\|^2 - \frac{1}{2}\|v_n\|^2\right) + E(u_{n+1}) - E(u_n) \le -Ch(\|v_{n+1}\|^2 + \|v_n\|^2)$$

has the same structure as (20) and guarantees that the adjusted total energy H^h is a strict Lyapunov function for the discrete dynamical system generated by (28). Furthermore, the dissipation of H^h is the same (up to constant C) as in the proof of Theorem 4.1 such that besides the convergence to equilibrium resulting from LaSalle's principle one also has exponential energy decay as in (16).

6. **Numerical experiments.**

6.1. **A one-dimensional double-well potential: eventual exponential decay and energy dissipation.** We start by considering the dissipative system (1) with a polynomial double-well potential for which the AVF integral can be computed explicitly $E : \mathbb{R} \to \mathbb{R}$, $E(u) = \frac{1}{4}\left(u^2 - 1\right)^2$. The (eventual) exponential decay of the potential energy $E(u_n)$ is shown in Fig. 1. When starting the evolution from a state $(u_0, v_0) = (0.1, 0)$ which is close to the unstable equilibrium point $(u^*, v^*) = (0, 0)$, at first, the potential energy changes slowly, until the system is accelerated towards the attracting equilibrium of minimal potential energy $(u^*, v^*) = (1, 0)$. This is a

worst-case scenario since, whenever the initial state lies in a convexity domain of the total energy, an accelerated convergence to equilibrium will occur without delay.

Although the qualitative behavior of both energy decay and state-space trajectories remains similar for a wide range of step size values, a closer inspection reveals that the onset of exponential decay changes with h. This is related to the fact that larger step sizes introduce more (numerical) dissipation, albeit in favorable way as far as minimization of E is concerned.
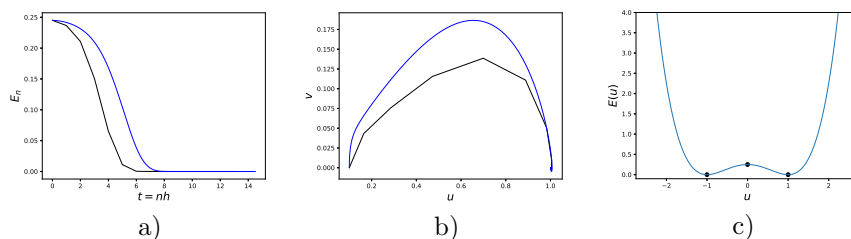

a) b) c)

FIGURE 1. The (eventual) exponential decay of the potential energy (a) as well as state-spaces trajectories (b) for decreasing step sizes $h = 1$ and 0.1, both with identical initial conditions $u_0 = 0.01, v_0 = 0$. The double-well potential is depicted in c).

6.2. **Nonconvex two-demensional potentials.** In order to gain perspective, we compare the proposed AVF splitting algorithm with the Störmer-Verlet conformal symplectic integrator (28) proposed in [5]. This explicit method is a viable less expensive alternative to the AVF only at small step sizes, as it becomes numerically unstable for large h. At the other end of the spectrum, for large step sizes, the structure-preserving AVF splitting is compared to a standard implicit scheme, the classical trapezoidal rule applied to the dissipative system (1)

$$\begin{cases} u_{n+1} = u_n + \dfrac{h}{2}\left(v_n + v_{n+1}\right) \\ v_{n+1} = v_n - h\left(v_n + v_{n+1}\right) - \dfrac{h}{2}\left(\nabla E\left(u_n\right) + \nabla E\left(u_{n+1}\right)\right) . \end{cases} \tag{32}$$

A nonconvex potential satisfying the Łojasiewicz inequality with $\theta = 1/2$. The first two-dimensional example that we consider is a nontrivial potential satisfying the Łojasiewicz inequality with coefficient $\theta = 1/2$ (see Ehrhardt et. al. [11]) in a small enough neighborhood of its unique minimum $u^* = (0,0)$, namely

$$E(u_1, u_2) = \frac{u_1^2}{1.5} + \frac{u_2^2}{1.5} + 3\sin^2\left(\frac{u_1 + u_2}{\sqrt{2}}\right). \tag{33}$$

Another feature of this potential is that it possesses two large symmetric, almost flat, plateaux away from its sole critical point (see Fig. 2 c)). We investigate the effect of such a plateau.

For a large step size $h = 1$ (see Fig. 2 a)), we compare the AVF splitting method with the implicit trapezoidal method (the Sörmer-Verlet conformal symplectic method being unstable at this step size). Both methods are only vaguely reproducing the energy plateau. The proposed algorithm, however manages to

detect its existence. Nevertheless, both methods do well in reproducing the asymptotic convergence to equilibrium. A default Dormand-Prince method with step size $h = 0.1$ (red) is used as a benchmark.

At a much smaller step size $h = 0.1$ (Fig. 2 b)) one can faithfully reproduce also the intermediate behavior (energy plateau), not just the asymptotic behavior. Here we compare the proposed AVF-based method to the conformal symplectic method of [5]. Both methods behave in a very similar manner and, in this case, the explicit method is a viable, less expensive alternative for the AVF splitting.

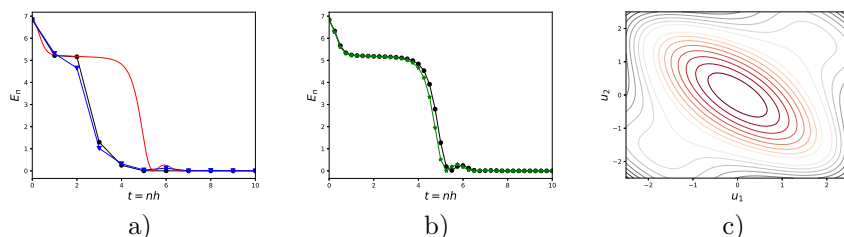a)   b)   c)

FIGURE 2. Capturing an energy plateau: a) the AVF Splitting algorithm (black) compared to the trapezoidal rule (blue) for $h = 1$ (and benchmark Runge-Kutta (red)); b) AVF Splitting algorithm (black) compared to the conformal symplectic algorithm (28) (green) for $h = 0.1$. The contour lines of the nonconvex potential (33) are depicted in c).

The two-dimensional Rosenbrock function: a potential exhibiting a narrow, flat valley. The new experiment is concerned with the two-dimensional Rosenbrock potential

$$E : \mathbb{R}^2 \to \mathbb{R}^2, \quad E(u_1, u_2) = (1 - u_1)^2 + 100(u_2 - u_1^2)^2.$$

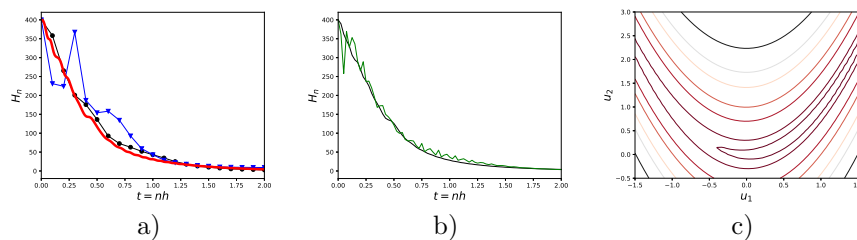a)   b)   c)

FIGURE 3. Erroneous total energy oscillations: a) AVF Splitting algorithm (black) compared to the trapezoidal rule (blue) for $h = 0.1$ (and benchmark Runge-Kutta (red)); b) AVF Splitting algorithm (black) compared to the conformal symplectic algorithm (28) (green) for $h = 0.01$. The contour lines of the Rosenbrock potential are depicted in c).

The global minimum $(u_1^*, u_2^*) = (1, 1)$ of this well-known performance test function for optimization lies inside a long and narrow, parabolic shaped flat valley. This function is analytic and hence Łojasiewicz arround its unique minimum, but

not with coefficient $\theta = 1/2$, such that our results apply except for the exponential rate of energy decay.

Throughout Section 2 we have emphasized the structural importance of the energy decrease for the continuous model. The fact that any discretization inherits this property cannot be taken for granted, as many well-established algorithms may fail in this respect. Indeed, one can observe in Fig. 3 that the proposed algorithm correctly reproduces the decrease of the total energy, while both the trapezoidal rule (32) and the conformal symplectic method (28) exhibit erroneous oscillations.

7. **Conclusions.** In the context of geometric numerical integration, Hamiltonian system have taken center for a long time, but the last couple of decades have witnessed growing interest in the analysis of damped systems. Beyond the obvious mechanical motivation, there is an additional dimension of dissipative systems pertaining to optimization and its applications such as Machine Learning.

The present work, deals with a dissipative Hamiltonian system, for which a splitting-based structure-preserving numerical algorithm that faithfully replicates the energy-dissipation properties of the continuous system is analyzed. The non-linear Hamiltonian subproblem is integrated using an energy-conserving averaged vector field (AVF) method that allows for exact replication of the continuous system's equilibrium points and, more importantly, of its strict Lyapunov function (the total energy). As a consequence of this deep structural similarity, the asymptotic convergence to equilibrium of the continuous system is recovered at a discrete level, with the same energy decay rate. The proof of this result does not rely on backward error analysis but rather on the aforementioned structural similarity. On the other hand, a C^2 regularity requirement is essential and we do not see any obvious way of dealing with less regular potentials in the same framework.

The high computational cost of the implicit AVF method is well-known and, whenever possible, one would prefer to replace it by an explicit (e.g., Störmer-Verlet) alternative. However, numerical experiments as well as the linear analysis in [5] show that the stability of such explicit methods comes at a high price in terms of step size restrictions and erroneous oscillations in total energy. Extending the present asymptotic analysis to methods based on explicit symplectic integrators with adaptive step size, as well as a detailed discussion of the small/vanishing dissipation behavior are possible future directions that we consider.

This study is essentialy nonconvex in nature so it does complement rather than compete with recent developments in dynamical convex optimization (see [3], [39], [6], [26], to mention just a few). Ever since Su, Boyd and Candès [41] have pointed out that

$$\ddot{u} + \frac{3}{t}\dot{u} = -\nabla E\left(u\right) \tag{34}$$

is a continuous model for Nesterov's accelerated gradient method, the topic has attracted a great deal of interest in the optimization community (cf. [3]), and an increasing interest in connections to numerical analysis (see [9] and [24]). The dynamical system generated by (34) is nonautonomous, nevertheless splitting leads to the separation of the linear, exactly solvable dissipative subsystem from the Hamiltonian one, exactly as in this contribution. Second-order symplectic splitting discretizations are expected to achieve acceleration in a similar way to first order ones introduced in [39], only with higher accuracy.

Form an optimization perspective, it is quite remarkable that both at the semidis-crete level and at the level of the full discretization energy decay and equilibrium

points are exactly preserved by AVF splitting and that, in contrast to common optimization algorithms (see [33]), there is no restriction on the step size which involves the Lipschitz constant of the objective's function gradient. This however, is valid only assuming exact solvability of the implicit problem. Up to our knowledge, the trade-off between accuracy in solving the implicit equation and step size is not fully understood, from a long-time asymptotics perspective.

Another natural question that arises concerns possible extensions of the finite-dimensional analysis presented here to an abstract infinite-dimensional setting. That is, to complement results concerning semilinear evolution equations in [19] or [12] with an asymptotic analysis of the discretizations behavior, provided that solutions of the continuous equation converge to equilibrium. For example, one can think of semilinear damped wave equations, which have the same dissipative structure as the one considered in this work. In [35], it has been dealt with such equations by combining Strang splitting with a finite element spatial discretization, except that the long-time regime has not been analyzed yet.

Appendix A. Omitted Proof from Section 4. The aim of this section is to present a detailed derivation of the inequality (26), that is

$$\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \leq -\frac{\varepsilon h}{2}\left(\widetilde{H}_n^{n+1} + \widetilde{H}_{n-1}^n\right).$$

To this end, we are still adhering to the notations (23), (24) and hypotheses of Theorem 4.1, while, for simplicity, we take $E(u^*) = 0$. Furthermore, $n \geq n_0$ with n_0 large enough such that $(u_n)_{n \geq n_0}$ lies within a neighborhood of u^* in which the Łojasiewicz inequality holds, that is

$$E_n \leq \frac{1}{2}\|\nabla E(u_n)\|^2. \tag{35}$$

By the definition (25) the right hand side of (26) can be written down in full detail as

$$
\begin{aligned}
\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \quad \leq \quad & -\frac{\varepsilon h}{2}(1+2h^2)\frac{\|v_{n+1}\|^2 + 2\|v_n\|^2 + \|v_{n-1}\|^2}{4} \\
& -\frac{\varepsilon h}{2}\frac{E_{n+1} + 2E_n + E_{n-1}}{2} \\
& -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla} E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N} \\
& -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla} E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N}.
\end{aligned}
\tag{36}
$$

In order to to arrive at this inequality, our strategy is to start with direct computations based on the definition of the modified energy $\widetilde{H}_n^{n+1}$ and the energy dissipation relation (20) such that in

$$
\begin{aligned}
\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \quad \leq \quad & -\frac{h}{2}D + \frac{\varepsilon}{2}\left\langle \overline{\nabla} E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N} \\
& -\frac{\varepsilon}{2}\left\langle \overline{\nabla} E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N},
\end{aligned}
\tag{37}
$$

we can construct all the desired terms at the cost of $O(\varepsilon)$ error terms which are dominated by the good dissipation term

$$-\frac{h}{2}D = -\frac{h}{2}(1+2h^2)\left(\|v_{n+1}\|^2 + 2\|v_n\|^2 + \|v_{n-1}\|^2\right).$$

We start by rewriting the scalar products (37) such that velocity differences are created

$$\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n = -\frac{h}{2}D$$
$$+ \frac{\varepsilon}{2} \left\langle \overline{\nabla}E_n^{n+1},\ e^h v_{n+1} - e^h v_n \right\rangle_{\mathbb{R}^N} + \frac{\varepsilon}{2} \left\langle \overline{\nabla}E_{n-1}^n,\ e^{-h} v_n - e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N}$$
$$+ \frac{\varepsilon}{2} \left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ e^h v_n \right\rangle_{\mathbb{R}^N} + \frac{\varepsilon}{2} \left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ e^{-h} v_n \right\rangle_{\mathbb{R}^N},$$

since these appear in the somewhat more lucrative form of the velocity equation in (15), that reminds of standard finite differences but with a reminder term $R_n^{n+1} = (e^h - 1 - h)v_{n+1} - (e^{-h} - 1 + h)v_n$

$$v_{n+1} - v_n = -h(v_{n+1} - v_n) - h\overline{\nabla}E_n^{n+1} - R_n^{n+1}. \tag{38}$$

Now, it follows from (38) that

$$\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \leq$$
$$-\frac{h}{2}D - \frac{\varepsilon h}{2} \left\langle \overline{\nabla}E_n^{n+1},\ e^h v_n - e^h v_{n+1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2} e^h \left\| \overline{\nabla}E_n^{n+1} \right\|^2$$
$$-\frac{\varepsilon}{2} e^h \left\langle \overline{\nabla}E_n^{n+1},\ R_n^{n+1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2} \left\langle \overline{\nabla}E_{n-1}^n,\ e^{-h} v_{n-1} - e^{-h} v_n \right\rangle_{\mathbb{R}^N}$$
$$-\frac{\varepsilon h}{2} e^{-h} \left\| \overline{\nabla}E_{n-1}^n \right\|^2 - \frac{\varepsilon}{2} e^{-h} \left\langle \overline{\nabla}E_{n-1}^n,\ R_{n-1}^n \right\rangle_{\mathbb{R}^N}$$
$$+\frac{\varepsilon}{2} \left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ (e^h + e^{-h})v_n \right\rangle_{\mathbb{R}^N}.$$

Bearing in mind that our aim is to reach (26), we construct scalar products containing $e^h v_{n+1} + e^{-h} v_n$ and group the terms as follows

$$\widetilde{H}_n^{n+1} - \widetilde{H}_{n-1}^n \leq -\frac{h}{2}D + T_{1a} + T_{1b} + T_2 + T_3,$$

$$T_{1a} = -\frac{\varepsilon h}{2} \left\langle \overline{\nabla}E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2} e^h \left\| \overline{\nabla}E_n^{n+1} \right\|^2$$
$$T_{1b} = -\frac{\varepsilon h}{2} \left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2} e^{-h} \left\| \overline{\nabla}E_{n-1}^n \right\|^2$$
$$T_2 = \frac{\varepsilon}{2} \left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ (e^h + e^{-h})v_n \right\rangle_{\mathbb{R}^N}$$
$$\quad - \frac{\varepsilon h}{2} \left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ (e^{-h} - e^h)v_n \right\rangle_{\mathbb{R}^N}$$
$$T_3 = -\frac{\varepsilon}{2} e^h \left\langle \overline{\nabla}E_n^{n+1},\ R_n^{n+1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon}{2} e^{-h} \left\langle \overline{\nabla}E_{n-1}^n,\ R_{n-1}^n \right\rangle_{\mathbb{R}^N}.$$

The reminder of the proof concentrates on finding appropriate upper bounds for each of these terms. In view of (36) we force a desired coefficient for the scalar product term in T_{1a}, that is

$$T_{1a} = -\frac{\varepsilon^2 h}{4} \left\langle \overline{\nabla}E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N}$$
$$\quad - \frac{\varepsilon h}{2}\left(1 - \frac{\varepsilon}{2}\right) \left\langle \overline{\nabla}E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2} c \left\| \overline{\nabla}E_n^{n+1} \right\|^2,$$

and proceed by using $e^{-h}v_n = e^h v_{n+1} + h\overline{\nabla}E_n^{n+1}$ together with the inequality $-ab \leq \frac{a^2}{4} + b^2$ to get

$$-\frac{\varepsilon h}{2}\left(1 - \frac{\varepsilon}{2}\right)\left\langle \overline{\nabla}E_n^{n+1}, \ e^h v_{n+1} + e^{-h}v_n \right\rangle_{\mathbb{R}^N} \leq$$

$$\frac{\varepsilon h e^h}{4}\left\|\overline{\nabla}E_n^{n+1}\right\|^2 + \varepsilon h e^h\left(1 - \frac{\varepsilon}{2}\right)^2\|v_{n+1}\|^2.$$

Based on the fact that $e^h > 1$,

$$T_{1a} \leq -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_n^{n+1}, \ e^h v_{n+1} + e^{-h}v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{4}\left\|\overline{\nabla}E_n^{n+1}\right\|^2$$

$$+ \varepsilon h e^h\left(1 - \frac{\varepsilon}{2}\right)^2\|v_{n+1}\|^2$$

holds true and, as in the analysis of the continuous case, we would like to use the Łojasiewicz inequality to produce potential energy terms on the right hand side of the previous inequality. This is not directly possible, since the term appearing contains the average not exact gradient. Nevertheless, after some preparations and at the cost of some controllable error terms, energy contributions can be produced. To this end, we start by observing that

$$\begin{aligned}
\|\nabla E(u_n)\|^2 &= \left\langle \nabla E(u_n), \ \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_n^{n+1} + \nabla E(u_n) \right\rangle_{\mathbb{R}^N} \\
&= \left\langle \nabla E(u_n), \ \overline{\nabla}E_n^{n+1} \right\rangle_{\mathbb{R}^N} \\
&\quad + \left\langle \overline{\nabla}E_n^{n+1} + \nabla E(u_n) - \overline{\nabla}E_n^{n+1}, \ \nabla E(u_n) - \overline{\nabla}E_n^{n+1} \right\rangle_{\mathbb{R}^N} \\
&\leq \frac{1}{2}\|\nabla E(u_n)\|^2 + \frac{1}{2}\left\|\overline{\nabla}E_n^{n+1}\right\|^2 \\
&\quad + \left\langle \overline{\nabla}E_n^{n+1}, \ \nabla E(u_n) - \overline{\nabla}E_n^{n+1} \right\rangle_{\mathbb{R}^N} + \left\|\nabla E(u_n) - \overline{\nabla}E_n^{n+1}\right\|^2,
\end{aligned}$$

or, rearranging,

$$\begin{aligned}
-\frac{1}{2}\left\|\overline{\nabla}E_n^{n+1}\right\|^2 &\leq -\frac{1}{2}\|\nabla E(u_n)\|^2 + \left\|\nabla E(u_n) - \overline{\nabla}E_n^{n+1}\right\|^2 \\
&\quad + \left\langle \overline{\nabla}E_n^{n+1}, \ \nabla E(u_n) - \overline{\nabla}E_n^{n+1} \right\rangle_{\mathbb{R}^N}.
\end{aligned} \tag{39}$$

In order to estimate both undesired terms we apply the mean value Theorem

$$\begin{aligned}
\nabla E(u_n) - \overline{\nabla}E_n^{n+1} &= \int_0^1 \left(\nabla E(u_n) - \nabla E((1-\alpha)u_n + \alpha u_{n-1})\right)d\alpha \\
&= -\left(\int_0^1 \nabla^2 E(\widetilde{u}_n^\alpha)\alpha\,d\alpha\right)(u_{n+1} - u_n) \\
&= -\left(\int_0^1 \nabla^2 E(\widetilde{u}_n^\alpha)\alpha\,d\alpha\right)\frac{h}{2}(e^h v_{n+1} + e^{-h}v_n),
\end{aligned}$$

Based on the fact that E is C^2 and since the sequence (u_n) remains in closed bounded domain, there exists $M > 0$ such that

$$\left\|\nabla E(u_n) - \overline{\nabla}E_n^{n+1}\right\|^2 \leq \frac{h^2}{4}M^2 C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right).$$

On the other hand, inserting $\overline{\nabla}E_n^{n+1} = -\frac{1}{h}(e^h v_{n+1} - e^{-h}v_n)$ in the scalar product term of (39) yields

$$\left|\left\langle \overline{\nabla}E_n^{n+1}, \ \nabla E(u_n) - \overline{\nabla}E_n^{n+1} \right\rangle_{\mathbb{R}^N}\right| \leq \frac{1}{4}MC\left(\|v_{n+1}\|^2 + \|v_n\|^2\right)$$

such that (39) becomes

$$-\frac{1}{2}\left\|\overline{\nabla}E_n^{n+1}\right\|^2 \leq -\frac{1}{2}\left\|\nabla E(u_n)\right\|^2 + C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right).$$

In a very similar manner, one deduces also the symmetric estimate

$$-\frac{1}{2}\left\|\overline{\nabla}E_n^{n+1}\right\|^2 \leq -\frac{1}{2}\left\|\nabla E(u_{n+1})\right\|^2 + C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right).$$

By taking the average of the two and using the Łojasiewicz inequality (35)

$$\begin{aligned}
-\left\|\overline{\nabla}E_n^{n+1}\right\|^2 &\leq -\frac{1}{2}\left(\|\nabla E(u_{n+1})\|^2 + \|\nabla E(u_{n+1})\|^2\right) + C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right) \\
&\leq -(E_{n+1} + E_n) + C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right).
\end{aligned}$$

Reverting to the estimate for T_{1a}, we finally have

$$\begin{aligned}
T_{1a} \leq\ & -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_n^{n+1},\ e^h v_{n+1} + e^{-h} v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2}\left(\frac{E_{n+1} + E_n}{2}\right) \\
& + \frac{\varepsilon h^3}{16} C\left(\|v_{n+1}\|^2 + \|v_n\|^2\right) + \varepsilon h e^h \|v_{n+1}\|^2.
\end{aligned} \tag{40}$$

When dealing with the term T_{1b} we take a similar approach. However, it is not a completely straightforward matter due to the asymmetry $e^h > 1$, $e^{-h} < 1$. Again, we start by forcing a desired scalar product term and deduce, from $e^{-h}v_{n-1} = e^h v_n + h\overline{\nabla}E_{n-1}^n$ that

$$\begin{aligned}
T_{1b} =\ & -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h^{-h}}{2}\left\|\overline{\nabla}E_{n-1}^n\right\|^2 \\
& - \frac{\varepsilon h}{2}\left(1 - \frac{\varepsilon}{2}\right)\left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N} \\
\leq\ & -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2}e^{-h}\left\|\overline{\nabla}E_{n-1}^n\right\|^2 \\
& - \frac{\varepsilon h}{2}\left(1 - \frac{\varepsilon}{2}\right)\left\langle \overline{\nabla}E_{n-1}^n,\ 2e^h v_n \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h^2}{2}\left(1 - \frac{\varepsilon}{2}\right)\left\|\overline{\nabla}E_{n-1}^n\right\|^2.
\end{aligned}$$

Then, choosing $\varepsilon < 1$, which assures $1 - \frac{\varepsilon}{2} > \frac{1}{2}$, and based on a weighted version of Young's inequality

$$\begin{aligned}
T_{1b} \leq\ & -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{4}\left\|\overline{\nabla}E_{n-1}^n\right\|^2 \\
& + \varepsilon h e^{3h}\left(1 - \frac{\varepsilon}{2}\right)^2 \|v_n\|^2.
\end{aligned}$$

Again, as in the case of T_{1a}, owing to the Łojasiewicz inequality

$$\begin{aligned}
T_{1b} \leq\ & -\frac{\varepsilon^2 h}{4}\left\langle \overline{\nabla}E_{n-1}^n,\ e^h v_n + e^{-h} v_{n-1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon h}{2}\left(\frac{E_n + E_{n-1}}{2}\right) \\
& + \frac{\varepsilon h^3}{16} C\left(\|v_n\|^2 + \|v_{n-1}\|^2\right) + \varepsilon h e^{3h}\|v_n\|^2.
\end{aligned} \tag{41}$$

Now, we deal with T_2. Introducing the notation $K(h) = (e^h + e^{-h}) - h(e^h - e^{-h})$ we can write $T_2 = K(h)\frac{\varepsilon}{2}\left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ v_n \right\rangle_{\mathbb{R}^N}$ and see that this term corresponds to the Hessian term $\nabla^2 E$ in the analysis of the continuous model. Indeed, for each $\alpha \in [0,1]$, the mean value theorem gives

$$\left\langle \overline{\nabla}E_n^{n+1} - \overline{\nabla}E_{n-1}^n,\ v_n \right\rangle_{\mathbb{R}^N} \leq \int_0^1 \left\|\nabla^2 E(\widetilde{u}_n^\alpha)\right\| \|\widetilde{v}_n^\alpha\| \|v_n\|\, \mathrm{d}\alpha$$

for $\widetilde{v}_n^\alpha = (1 - \alpha)(u_n - u_{n-1}) + \alpha(u_{n+1} - u_n)$ and intermediate points $\widetilde{u}_n^\alpha$ on the segment $[(1-\alpha)u_{n-1} + \alpha u_n, (1-\alpha)u_n + \alpha u_{n+1}]$.

Here, since for all $n \geq n_0$, u_n lies in a closed, bounded, small enough energy sublevel set of E and E is C^2, there exists a constant $M > 0$, such that $\left\| \nabla^2 E(\widetilde{u}_n^\alpha) \right\|_{Op} \leq M$ for all $\alpha \in [0,1]$ and $n \geq n_0$. Also, as $\alpha \in [0,1]$, we have

$$\|\widetilde{v}_n^\alpha\| \leq \|u_{n+1} - u_n\| + \|u_n - u_{n-1}\| = \frac{h}{2} \left(e^h \|v_{n+1}\| + (e^h + e^{-h})\|v_n\| + e^{-h}\|v_{n-1}\| \right).$$

Combining all these yields, for some positive constant $C > 0$,

$$T_2 \leq \frac{\varepsilon h}{4} C \left(\|v_{n+1}\|^2 + 2\|v_n\|^2 + \|v_{n-1}\|^2 \right). \tag{42}$$

The final term $T_3 = -\frac{\varepsilon}{2} e^h \left\langle \overline{\nabla} E_n^{n+1}, \ R_n^{n+1} \right\rangle_{\mathbb{R}^N} - \frac{\varepsilon}{2} e^{-h} \left\langle \overline{\nabla} E_{n-1}^n, \ R_{n-1}^n \right\rangle_{\mathbb{R}^N}$ can also be estimated in terms of square norms of consecutive velocities by recalling that $R_n^{n+1} = (e^h - 1 - h)v_{n+1} - (e^{-h} - 1 + h)v_n$ together with

$$-\overline{\nabla} E_n^{n+1} = \frac{1}{h}(e^h v_{n+1} - e^{-h} v_n) \quad \text{and} \quad -\overline{\nabla} E_{n-1}^n = \frac{1}{h}(e^h v_n - e^{-h} v_{n-1}).$$

Actually, one has

$$T_3 \leq \frac{\varepsilon h}{2} C \left(\|v_{n+1}\|^2 + 2\|v_n\|^2 + \|v_{n-1}\|^2 \right), \tag{43}$$

for some positive constant C.

Taking into account all the estimates (40),(41),(42) and (43) for the four terms T_{1a}, T_{1b}, T_2 and T_3 we can observe that the desired inequality (36) has been obtained since, for ε is small enough, the dissipation term $-\frac{h}{2} D$ dominates the error terms that we have created and which all have $O(\varepsilon)$ coefficients.

Appendix B. Omitted estimates for the Störmer-Verlet splitting. This section deals in more detail with the two reminder terms on the right hand side of (30). More precisely, we show that both

$$\Phi_1 = \frac{h^2}{8} \left\langle \nabla E(u_n) + \nabla E(u_{n+1}), \ \nabla E(u_{n+1}) - \nabla E(u_n) \right\rangle_{\mathbb{R}^N}, \tag{44}$$

$$\Phi_2 = \frac{h^2}{8} \left\langle \overline{\nabla} E_n^{n+1} - \frac{\nabla E(u_n) + \nabla E(u_{n+1})}{2}, \ u_{n+1} - u_n \right\rangle_{\mathbb{R}^N} \tag{45}$$

are bounded from above by velocity terms.

The mean value theorem allows us to rewrite the first equation of (29) as

$$u_{n+1} - u_n = \frac{h}{2}(e^h v_{n+1} + e^{-h} v_n) + \frac{h^2}{4} \nabla^2 E(\xi_n^{n+1})(u)_{n+1} - u_n) \quad \text{for} \quad \xi_n^{n+1} \in [u_n, u_{n+1}]$$

and in view of the C^2 regularity and the Lipschitz condition for ∇E, $\left\| \nabla^2 E(\xi_n^{n+1}) \right\| < L$ the step size conditions (31) implies $\frac{h^2}{4} \left\| \nabla^2 E(\xi_n^{n+1}) \right\| < 1$. So

$$u_{n+1} - u_n = \left(I - \frac{h^2}{4} \nabla^2 E(\xi_n^{n+1}) \right)^{-1} \frac{h}{2}(e^h v_{n+1} + e^{-h} v_n)$$

an finally

$$\|u_{n+1} - u_n\| \leq \frac{1}{1 - \frac{h^2}{4} L} \frac{h}{2} \left\| e^h v_{n+1} + e^{-h} v_n \right\|. \tag{46}$$

Now, we are in position to estimate Φ_1 and Φ_2. On one hand, based on the first equation in (28) and the convexity of E

$$
\begin{aligned}
\Phi_1 &\leq \frac{h^2}{8} \left\| \nabla E(u_{n+1}) - \nabla E(u_n) \right\|^2 + \frac{h^2}{4} \left\langle \nabla E(u_n),\ \nabla E(u_{n+1}) - \nabla E(u_n) \right\rangle_{\mathbb{R}^N} \\
&\leq \frac{h^2}{8} L^2 \left(\frac{1}{1 - \frac{h^2}{4}L} \right)^2 \frac{h^2}{4} \left\| e^h v_{n+1} + e^{-h} v_n \right\|^2 \\
&\quad + \frac{h e^{-h}}{2} \left\| v_n \right\| \frac{L}{1 - \frac{h^2}{4}L} \frac{h}{2} \left\| e^h v_{n+1} + e^{-h} v_n \right\|
\end{aligned}
$$

which leads, for h small enough, to the desired estimate, i.e., $\Phi_1 \leq Ch(\|v_{n+1}\|^2 + \|v_n\|^2)$.

On the other hand, to deal with Φ_2 we observe that

$$
\begin{aligned}
2\overline{\nabla} E_n^{n+1} - \nabla E(u_n) - \nabla E(u_{n+1}) &= \int_0^1 \nabla E\left((1-\alpha)u_n + \alpha u_{n+1} \right) - \nabla E(u_n)\, d\alpha \\
&\quad + \int_0^1 \nabla E\left((1-\alpha)u_n + \alpha u_{n+1} \right) - \nabla E(u_{n+1})\, d\alpha
\end{aligned}
$$

and the mean value theorem together with $((1-\alpha)u_n + \alpha u_{n+1}) - u_n = \alpha(u_{n+1} - u_n)$, $((1-\alpha)u_n + \alpha u_{n+1}) - u_{n+1} = (\alpha - 1)(u_{n+1} - u_n)$ and (46) leads to the desired conclusion

$$
\begin{aligned}
\Phi_2 &= \frac{1}{2} \left\langle \left[\int_0^1 \left(\alpha \nabla^2 E(\eta(\alpha)) + (\alpha - 1)\nabla^2 E(\zeta(\alpha)) \right) d\alpha \right] \right. \\
&\quad \left. (u_{n+1} - u_n),\ (u_{n+1} - u_n) \right\rangle_{\mathbb{R}^N} \\
&\leq L \left\| u_{n+1} - u_n \right\|^2 \leq \frac{L}{1 - \frac{h^2}{4}L} h^2 \left\| e^h v_{n+1} + e^{-h} v_n \right\|^2.
\end{aligned}
$$

REFERENCES

[1] C. D. Alecsa, S. C. László and A. Viorel, A gradient-type algorithm with backward inertial steps associated to a nonconvex minimization problem, *Numer. Algorithms*, **84** (2020), 485–512.

[2] H. Attouch, X. Goudou and P. Redont, The heavy ball with friction method. I. The continuous dynamical system: Global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system, *Commun. Contemp. Math.*, **2** (2000), 1–34.

[3] H. Attouch, Z. Chbani, J. Peypouquet and P. Redont, Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity, *Math. Program.*, **168** (2018), 123–175.

[4] P. Bégout, J. Bolte and M. A. Jendoubi, On damped second-order gradient systems, *J. Differential Equations*, **259** (2015), 3115–3143.

[5] A. Bhatt, D. Floyd and B. E. Moore, Second order conformal symplectic schemes for damped Hamiltonian systems, *J. Sci. Comput.*, **66** (2016), 1234–1259.

[6] R. I. Boţ, E. R. Csetnek and S. C. László, Approaching nonsmooth nonconvex minimization through second-order proximal-gradient dynamical systems, *J. Evol. Equ.*, **18** (2018), 1291–1318.

[7] T. Cazenave and A. Haraux, *An Introduction to Semilinear Evolution Equations*, vol. 13 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press, Oxford University Press, New York, 1998, Translated from the 1990 French original by Yvan Martel and revised by the authors.

[8] E. Celledoni, R. I. McLachlan, D. I. McLaren, B. Owren, G. R. W. Quispel and W. M. Wright, Energy-preserving Runge-Kutta methods, *M2AN Math. Model. Numer. Anal.*, **43** (2009), 645–649.

[9] J. Diakonikolas and M. I. Jordan, Generalized momentum-based methods: A hamiltonian perspective, preprint, `arXiv:1906.00436`.

[10] G. Dujardin and P. Lafitte, Asymptotic behaviour of splitting schemes involving time-subcycling techniques, *IMA J. Numer. Anal.*, **36** (2016), 1804–1841.

[11] M. J. Ehrhardt, E. S. Riis, T. Ringholm and C.-B. Schönlieb, A geometric integration approach to smooth optimisation: Foundations of the discrete gradient method, preprint, `arXiv:1805.06444`.

[12] L. Einkemmer and A. Ostermann, Overcoming order reduction in diffusion-reaction splitting. Part 1: Dirichlet boundary conditions, *SIAM J. Sci. Comput.*, **37** (2015), A1577–A1592.

[13] E. Emmrich, Discrete versions of gronwall's lemma and their application to the numerical analysis of parabolic problems, *Preprint No. 637*, URL `https://www.math.uni-bielefeld.de/~emmrich/public/prepA.pdf`.

[14] G. França, J. Sulam, D. P. Robinson and R. Vidal, Conformal symplectic and relativistic optimization, preprint, `arXiv:1903.04100`.

[15] L. Gauckler, E. Hairer and C. Lubich, Dynamics, numerical analysis, and some geometry, in *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, World Sci. Publ., Hackensack, NJ, 2018, 453–485.

[16] E. Hairer, Energy-preserving variant of collocation methods, *JNAIAM. J. Numer. Anal. Ind. Appl. Math.*, **5** (2010), 73–84.

[17] E. Hairer and C. Lubich, Energy-diminishing integration of gradient systems, *IMA J. Numer. Anal.*, **34** (2014), 452–461.

[18] J. K. Hale, *Asymptotic Behavior of Dissipative Systems*, vol. 25 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1988.

[19] E. Hansen, F. Kramer and A. Ostermann, A second-order positivity preserving scheme for semilinear parabolic problems, *Appl. Numer. Math.*, **62** (2012), 1428–1435.

[20] A. Haraux and M. A. Jendoubi, Convergence of solutions of second-order gradient-like systems with analytic nonlinearities, *J. Differential Equations*, **144** (1998), 313–320.

[21] A. Haraux and M. A. Jendoubi, Decay estimates to equilibrium for some evolution equations with an analytic nonlinearity, *Asymptot. Anal.*, **26** (2001), 21–36.

[22] A. Haraux, *Systèmes Dynamiques Dissipatifs et Applications*, vol. 17 of Recherches en Mathématiques Appliquées [Research in Applied Mathematics], Masson, Paris, 1991.

[23] A. Haraux and M. A. Jendoubi, *The Convergence Problem for Dissipative Autonomous Systems, Classical Methods and Recent Advances*, BCAM SpringerBriefs. SpringerBriefs in Mathematics, Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2015,

[24] M. I. Jordan, Dynamical symplectic and stochastic perspectives on gradient-based optimization, in *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, World Sci. Publ., Hackensack, NJ, 2018, 523–549.

[25] J. P. LaSalle, *The Stability and Control of Discrete Processes*, With a foreword by Jack K. Hale and Kenneth R. Meyer, vol. 62 of Applied Mathematical Sciences, Springer-Verlag, New York, 1986.

[26] S. C. László, Convergence rates for an inertial algorithm of gradient type associated to a smooth nonconvex minimization, preprint, `arXiv:1811.09616`.

[27] S. Łojasiewicz, Une propriété topologique des sous-ensembles analytiques réels, in *Les Équations aux Dérivées Partielles (Paris, 1962)*, Éditions du Centre National de la Recherche Scientifique, Paris, 1963, 87–89.

[28] S. MacNamara and G. Strang, Operator splitting, in *Splitting Methods in Communication, Imaging, Science, and Engineering*, Sci. Comput., Springer, Cham, 2016, 95–114.

[29] C. J. Maddison, D. Paulin, Y. W. Teh, B. O'Donoghue and A. Doucet, Hamiltonian descent methods, preprint, `arXiv:1809.05042`.

[30] R. McLachlan and M. Perlmutter, Conformal Hamiltonian systems, *J. Geom. Phys.*, **39** (2001), 276–300.

[31] R. I. McLachlan and G. R. W. Quispel, Splitting methods, *Acta Numer.*, **11** (2002), 341–434.

[32] K. Modin and G. Söderlind, Geometric integration of Hamiltonian systems perturbed by Rayleigh damping, *BIT*, **51** (2011), 977–1007.

[33] Y. Nesterov, *Lectures on Convex Optimization*, Second edition of [MR2142598], vol. 137 of Springer Optimization and Its Applications, Springer, Cham, 2018.

[34] B. T. Polyak, Some methods of speeding up the convergence of iterative methods, *Ž. Vyčisl. Mat. i Mat. Fiz.*, **4** (1964), 791–803.

[35] A. Quaini and R. Glowinski, Splitting methods for some nonlinear wave problems, in *Splitting Methods in Communication, Imaging, Science, and Engineering*, Sci. Comput., Springer, Cham, 2016, 643–676.

[36] G. R. W. Quispel and D. I. McLaren, A new class of energy-preserving numerical integration methods, *J. Phys. A*, **41** (2008), 045206, 7 pp.

[37] I. Segal, Non-linear semi-groups, *Ann. of Math.*, **78** (1963), 339–364.

[38] X. Shang and H. C. Öttinger, Structure-preserving integrators for dissipative systems based on reversible–irreversible splitting, *Proceedings of the Royal Society A*, **476** (2020), 20190446, 25 pp.

[39] B. Shi, S. S. Du, W. Su and M. I. Jordan, Acceleration via symplectic discretization of high-resolution differential equations, in *Advances in Neural Information Processing Systems*, 2019, 5745–5753.

[40] G. Strang, On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.*, **5** (1968), 506–517.

[41] W. Su, S. Boyd and E. J. Candès, A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights, *J. Mach. Learn. Res.*, **17** (2016), Paper No. 153, 43 pp.

[42] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, vol. 68 of Applied Mathematical Sciences, 2nd edition, Springer-Verlag, New York, 1997.

[43] J. Zhang, A. Mokhtari, S. Sra and A. Jadbabaie, Direct runge-kutta discretization achieves acceleration, in *Advances in Neural Information Processing Systems*, 2018, 3900–3909.

E-mail address: adrian.viorel@math.ubbcluj.ro

E-mail address: alecsa@rist.ro

E-mail address: t.pinta@math.uni-goettingen.de