Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

The elliptic-regularization principle in Lagrangian mechanics

Matthias Liero[1], Ulisse Stefanelli[2]

submitted: November 18, 2011

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: matthias.liero@wias-berlin.de
http://www.wias-berlin.de/~liero/

[2] IMATI "E. Magenes" - CNR
v. Ferrata 1
27100 Pavia
Italy
E-Mail: ulisse.stefanelli@imati.cnr.it
http://www.imati.cnr.it/ulisse/

No. 1662
Berlin 2011

2010 *Mathematics Subject Classification.* 70H03, 70H30, 65L10.

Key words and phrases. Lagrangian mechanics, variational principle, elliptic regularization, time discretization.

U.S. and M.L. are partly supported by FP7-IDEAS-ERC-StG Grant # 200947 *BioSMA*. U.S. acknowledges the partial support of CNR-AVČR Grant *SmartMath*, and the Alexander von Humboldt Foundation. Furthermore, M.L. thanks the IMATI Pavia, where part of the work was conducted, for its kind hospitality.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We present a novel variational approach to Lagrangian mechanics based on elliptic regularization with respect to time. A class of parameter-dependent global-in-time minimization problems is presented and the convergence of the respective minimizers to the solution of the system of Lagrange's equations is ascertained. Moreover, we extend this perspective to mixed dissipative/nondissipative situations, present a finite time-horizon version of this approach, and provide related Γ-convergence results. Finally, some discussion on corresponding time-discrete versions of the principle is presented.

1 Introduction

Variational principles in Continuum Mechanics and Thermodynamics have been the subject of constant attention since their early appearance more than two centuries ago. From the philosophical viewpoint the investigation on variational principles is of a paramount importance for it corresponds to the fundamental quest for general and simple explanations of reality as we experience it. On the other hand, beside their indisputable elegance, variational principles have a clear practical impact as they originate a wealth of new perspectives and serve as unique tools for the analysis of real physical situations. Correspondingly, the mathematical literature on variational principles in Mechanics is overwhelming and a number of monographs on the subject are available. Being completely beyond our purposes to attempt a comprehensive review of the development of this subject, we shall minimally refer the reader to some classical monographs [Lán70, Moi04] as well as the more recent [Bas07, Ber09, Gho08].

The focus of this note is to present a new variational principle in the context of classical Lagrangian Mechanics. In particular, we shall be concerned with the evolution of a conservative dynamical system described by a set of generalized coordinates $\boldsymbol{q} \in \mathbb{R}^m$ $(m \in \mathbb{N})$ and characterized by the *Lagrangian* [Arn89]

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) := \frac{1}{2}\dot{\boldsymbol{q}}\cdot M\dot{\boldsymbol{q}} - U(\boldsymbol{q}).$$

Here, M is the symmetric and positive definite *mass matrix*, so that $\dot{\boldsymbol{q}}\cdot M\dot{\boldsymbol{q}}/2$ is the classical *kinetic energy* term. Moreover, we assume to be given the *potential energy* $U \in \mathrm{C}^{1,1}(\mathbb{R}^m)$ which we ask to be bounded from below.

Our aim is to present a novel variational view at Lagrangian mechanics based on elliptic regularization (ER) with respect to time. We shall be considering the minimization of the global-in-time functionals ER_ε defined on entire trajectories $\boldsymbol{q} : \mathbb{R}_+ \to \mathbb{R}^m$ as

$$\mathsf{ER}_\varepsilon[\boldsymbol{q}] := \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2}{2}\ddot{\boldsymbol{q}}(t)\cdot M\ddot{\boldsymbol{q}}(t) + U(\boldsymbol{q}(t))\right) \mathrm{d}t \qquad (\varepsilon > 0).$$

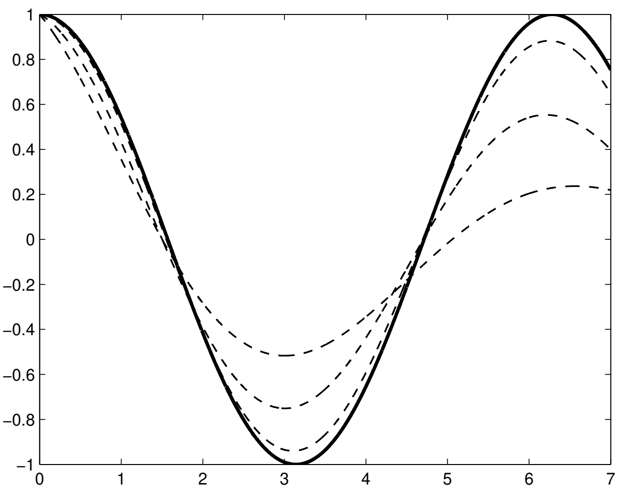

Figure 1: Convergence for $m = 1$, $U(q) = q^2/2$, $q^0 = 1$, and $q^1 = 0$. As $\varepsilon \to 0$, the minimizers of ER_ε on K_ε for $\varepsilon = 0.3, 0.1, 0.02$ (dashed) approach locally uniformly the solution of (1), namely $t \mapsto \cos t$ (solid).

Note that the small parameter ε above has the physical dimension of time, so that the whole integrand in ER_ε is an energy. The functional ER_ε admits minimizers $\boldsymbol{q}_\varepsilon$ in the closed and convex set

$$K_\varepsilon := \{M\dot{\boldsymbol{q}},\ \boldsymbol{q} \in \mathrm{H}^1(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)\ :\ \boldsymbol{q}(0) = \boldsymbol{q}^0,\ M\dot{\boldsymbol{q}}(0) = M\boldsymbol{q}^1\}$$

where given initial data $\boldsymbol{q}^0 \in \mathbb{R}^m$ and $M\boldsymbol{q}^1 \in \mathbb{R}^m$ are prescribed (see Lemma 2.1 below).

A first result of this paper asserts that the minimizers $\boldsymbol{q}_\varepsilon$ of the functional ER_ε on K_ε converge to the unique solution of the system of Lagrange's equations (*Lagrangian system* for short in the sequel). Namely, we have the following.

Theorem 1.1 (ER principle)**.** *Let $\boldsymbol{q}_\varepsilon$ minimize ER_ε in K_ε. Then, $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ locally uniformly where $\boldsymbol{q}$ is the classical solution of the Lagrangian system*

$$M\ddot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \ \text{ in } \mathbb{R}_+, \quad \boldsymbol{q}(0) = \boldsymbol{q}^0, \quad M\dot{\boldsymbol{q}}(0) = M\boldsymbol{q}^1. \tag{1}$$

In the easiest possible setting, namely the scalar $(m = 1)$ and linear case of $U(q) = q^2/2$ with $q^0 = 1$ and $q^1 = 0$ the convergence result of Theorem 1.1 is illustrated in Figure 1.

The ER principle provides a new variational reformulation of the Lagrangian system (1) as a (limit of a class of) constrained minimization problem(s). Although the Cauchy problem for the Lagrangian system (1) is quite standard and can be analyzed directly, the formulation of the ER principle paves the way to the treatment of the system by purely variational means. In particular, the ER formalism allows for the *direct* application of the tools of the Calculus of Variations (the Direct Method and Γ-convergence, for instance) to the evolutive differential system (1). Let us mention that the elliptic-regularization (ER) approach bear its name from the fact that the Euler-Lagrange equations for the functional ER_ε read (see Subsection 2.3 below)

$$\varepsilon^2 M\boldsymbol{q}^{(4)} - 2\varepsilon M\boldsymbol{q}^{(3)} + M\ddot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \qquad \text{in } \mathbb{R}_+, \tag{2}$$

where $\boldsymbol{q}^{(k)}$ stands for the k-th derivative. In particular, the minimizers $\boldsymbol{q}_\varepsilon$ solve a fourth-order elliptic-in-time regularization of the Lagrangian system (1).

The ER method appears to be rather general. In particular, besides the basic convergence result of Theorem 1.1, we aim here at showing its possible extension in two relevant directions. On the one hand, in Section 4 we focus on a specific finite-time horizon version of the ER principle where the integration is confined to some finite-time interval $(0,T)$. This allows to sharpen the convergence result in order to obtain error rates and turns out to be better suited for the purpose of numerical investigation. On the other hand, we extend the ER principle to treat mixed dissipative/nondissipative situations such that of viscous dynamics, both in the infinite (Section 3) and finite-time horizon (Subsection 4.5). Eventually, in Section 5 we provide some Γ-convergence analysis for the limiting purely dissipative (viscous) and purely nondissipative cases as well as for the connection between the finite-time and the infinite-time horizon situation.

We shall now turn to the illustration of some of the specific features of the ER principle by focusing on its comparison with the more classical *Hamilton principle*. The latter asserts that actual trajectories of the Lagrangian system (1) on the time interval $(0,T)$ are extremizers of the *action* functional

$$S[\boldsymbol{q}] = \int_0^T \left(\frac{1}{2}\dot{\boldsymbol{q}}(t)\cdot M\dot{\boldsymbol{q}}(t) - U(\boldsymbol{q}(t))\right)\mathrm{d}t$$

among all paths with prescribed initial and final states $\boldsymbol{q}^0$ and $\boldsymbol{q}^T$. In particular, the Lagrangian system (1) exactly corresponds to the Euler-Lagrange equations for S.

The ER principle differs from the Hamilton principle in *three* crucial ways. First, Hamilton's principle is indeed a *stationarity principle* for it generally corresponds to the quest for a saddle point of the action functional (note however that this will be a true minimum for small T). On the contrary, the ER principle relies on a true constrained *minimization*.

Secondly, the ER principle is directly formulated on the whole time semiline $\mathbb{R}_+$ whereas Hamilton's approach calls for the specification of an artificial finite-time interval $(0,T)$ and a final state. In particular, the ER principle directly encodes directionality of time by explicitly requiring the knowledge of just *initial states*.

Finally, the ER principle is clearly not invariant by time reversal. This is indeed crucial as the ER perspective is naturally incorporating dissipative effects (see Section 3) thus qualifying it as a suitable tool in order to discuss limiting mixed dissipative/nondissipative dynamics. Note that dissipative effects cannot be directly treated via Hamilton's framework and one resorts in considering the classical Lagrange-D'Alembert principle instead.

The price to pay within the ER frame with respect to Hamilton's is the check of the extra limit $\varepsilon \to 0$. This is exactly the object of Theorem 1.1 and the main concern of this theory.

1.1 Review of the literature on the ER principle

Our interest for the ER principle has been inspired by a conjecture by De Giorgi [Gio96] on hyperbolic evolution. In particular, in [Gio96] it is conjectured that the minimizers of the PDE version of ER$_\varepsilon$

$$u \mapsto \int_0^T \int_{\mathbb{R}^m} \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2}{2} |\partial_{tt} u(x,t)|^2 + \frac{1}{2} |\nabla u(x,t)|^2 + \frac{1}{p} |u(x,t)|^p \right) \mathrm{d}x \, \mathrm{d}t \quad (p > 2)$$

among all space-time functions u with prescribed initial conditions, converge as $\varepsilon \to 0$ to a solution of the semilinear wave equation

$$\partial_{tt} u - \nabla u + |u|^{p-2} u = 0 \text{ in } \mathbb{R}^m \times \mathbb{R}_+.$$

This conjecture has been checked positively for $T < \infty$ in [Ste11] and then for $T = \infty$ by SERRA & TILLI in [ST10]. Already in [Gio96, Rem. 1] it is speculated that some similar result could hold for more general functionals of the Calculus of Variations as well. Our main result Theorem 1.1 provides here a positive answer to this extension of the conjecture in the finite-dimensional case.

Before being applied to hyperbolic equations, the ER principle has been considered in a variety of parabolic situations. In the linear case, some results can be found in the classical monograph by LIONS & MAGENES [LM72]. As for the nonlinear case, this procedure has been followed by ILMANEN [Ilm94] for proving existence and partial regularity of the so-called Brakke mean curvature flow of varifolds. Two examples of relaxation of gradient flows related to microstructure evolution are provided by CONTI & ORTIZ [CO08]. In the finite-time framework, the case of gradient flows in Hilbert and metric spaces along with a number of related results, has been considered in [MS11] and [RSSS11b, RSSS11a], respectively. Moreover, the ER technique has been extended to rate-independent evolution by MIELKE & ORTIZ [MO08] and further detailed in [MS08], whereas the doubly nonlinear parabolic case is addressed in [AS10, AS11]. Finally, an application of the ER principle in the context of gradient flows driven by linear-growth functionals is given in [SS11].

2 Elliptic regularization in $\mathbb{R}_+$

We focus here on the infinite-time horizon result of Theorem 1.1. With no loss of generality, hereafter we shall assume the potential U to be nonnegative. Moreover, in order to avoid cumbersome notation we shall let $M = \rho I$ with $\rho > 0$ and I is the identity matrix from here on. It should be however clear that the corresponding proofs for a general positive-definite mass matrix M can be then obtained with no particular intricacy.

A *caveat* on notation: In the remainder of the paper c stands for any positive constant, possibly depending on $|\boldsymbol{q}^0|,\ |\boldsymbol{q}^1|$, $\|\nabla U\|_{\mathrm{C}^{1,1}}$, $U(\boldsymbol{q}^0)$, and $|\nabla U(\mathbf{0})|$, and changing from line to line. Note specifically that c does not depend on ρ and, later, ν and T.

2.1 Existence of minimizers

Let us firstly record that minimizers of ER$_\varepsilon$ on K_ε actually exist.

Lemma 2.1 (Direct method)**.** ER$_\varepsilon$ *admits a minimizer in* K_ε*.*

Proof. Every minimizing sequence $\boldsymbol{q}_k \in K_\varepsilon$ fulfills $\rho \int_0^\infty \mathrm{e}^{-t/\varepsilon}|\ddot{\boldsymbol{q}}_k|^2 \leq c$ and it is hence compact in $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$. Upon extracting some subsequence, one can exploit the lower semicontinuity of U and pass to the $\liminf$ by Fatou's Lemma. □

2.2 A priori estimate

The proof of Theorem 1.1 relies on an a priori estimate on the minimizers $\boldsymbol{q}_\varepsilon$ of ER_ε on K_ε. We have the following.

Lemma 2.2 (A priori estimate)**.** *Let $\boldsymbol{q}_\varepsilon$ minimize ER_ε on K_ε. Then,*

$$\rho|\dot{\boldsymbol{q}}_\varepsilon(t)|^2 \leq c \quad \forall t > 0. \tag{3}$$

The lemma follows from the argument by SERRA & TILLI in [ST10] where the PDE case of semilinear wave equations is treated. We hence claim no originality here. Still, we record the proof of Lemma 2.2 for the sake of later reference with respect to its extension to the mixed dissipative/nondissipative case presented in Subsection 3 below.

Proof. Assume $\boldsymbol{q}_\varepsilon$ to be a minimizer and rescale time by letting $\boldsymbol{p}(t) := \boldsymbol{q}_\varepsilon(\varepsilon t)$. We define the rescaled functional G_ε as

$$G_\varepsilon[\boldsymbol{p}] := \int_0^\infty \mathrm{e}^{-t}\left(\frac{\rho}{2}|\ddot{\boldsymbol{p}}(t)|^2 + \varepsilon^2 U(\boldsymbol{p}(t))\right)\mathrm{d}t$$

so that $\varepsilon\mathrm{ER}_\varepsilon[\boldsymbol{q}_\varepsilon] = G_\varepsilon[\boldsymbol{p}]$. At first, let us note that, by choosing $\widehat{\boldsymbol{p}}(t) := \boldsymbol{q}^0 + (\varepsilon\boldsymbol{q}^1)t$ (which, in particular, is such that $t \mapsto \widehat{\boldsymbol{q}}(t) := \widehat{\boldsymbol{p}}(t/\varepsilon) \in K_\varepsilon$) and using the Lipschitz continuity of U, one has

$$G_\varepsilon[\boldsymbol{p}] \leq G_\varepsilon[\widehat{\boldsymbol{p}}] = \varepsilon^2\int_0^\infty \mathrm{e}^{-t}U(\widehat{\boldsymbol{p}}(t))\mathrm{d}t \leq c\varepsilon^2\int_0^\infty \mathrm{e}^{-t}\left(1+\varepsilon|\boldsymbol{q}^1|t\right)\mathrm{d}t \leq c\varepsilon^2. \tag{4}$$

In the following, we shall make use of the following elementary inequality [ST10, Lemma 2.3]

$$\int_t^\infty \mathrm{e}^{-s}f^2(s)\mathrm{d}s \leq 2\mathrm{e}^{-t}f^2(t) + 4\int_t^\infty \mathrm{e}^{-s}\dot{f}^2(s)\mathrm{d}s \tag{5}$$

which follows by integration by parts and is valid for all $f \in \mathrm{H}^1_{\mathrm{loc}}(\mathbb{R}_+)$ and $t \geq 0$, regardless of the finiteness of the integrals. In particular, we exploit inequality (5) in order to get that

$$\int_0^\infty \mathrm{e}^{-s}|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s \leq 2\varepsilon^2|\boldsymbol{q}^1|^2 + 4\int_0^\infty \mathrm{e}^{-s}|\ddot{\boldsymbol{p}}(s)|^2\mathrm{d}s \leq c\varepsilon^2 + \frac{c}{\rho}G_\varepsilon[\boldsymbol{p}]. \tag{6}$$

The latter entails that $t \mapsto \mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2 \in \mathrm{W}^{1,1}(\mathbb{R}_+)$ so that $\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2 \to 0$ as $t \to \infty$.

Define now, for all $t \geq 0$, the auxiliary function

$$H(t) := \int_t^\infty \mathrm{e}^{-s}\left(\frac{\rho}{2}|\ddot{\boldsymbol{p}}(s)|^2 + \varepsilon^2 U(\boldsymbol{p}(s))\right)\mathrm{d}s$$

and note that $H \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}_+)$, it is nonincreasing and nonnegative.

By considering competitors $\tilde{\boldsymbol{p}}(t) = \boldsymbol{p}(s(t))$ where s is some smooth time reparametrization, the minimality of $\boldsymbol{p}$ and the computations in [ST10, Prop. 3.1] ensure that

$$\left(\frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}}\right)^{\cdot} = \frac{1}{2}\left(\mathrm{e}^t H(t)\right)^{\cdot} + \rho|\ddot{\boldsymbol{p}}|^2 + \frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}}. \tag{7}$$

Let a second auxiliary function E be defined as

$$E(t) := \frac{\rho}{4}|\dot{\boldsymbol{p}}(t)|^2 - \frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}} + \frac{1}{2}\mathrm{e}^t H(t).$$

By virtue of relation (7) we compute that

$$\begin{aligned}\dot{E} &= \frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}} - \frac{\rho}{2}(\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}})^{\cdot} + \frac{1}{2}(\mathrm{e}^t H(t))^{\cdot} \\ &\overset{(7)}{=} \frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}} - \left(\frac{1}{2}\left(\mathrm{e}^t H(t)\right)^{\cdot} + \rho|\ddot{\boldsymbol{p}}|^2 + \frac{\rho}{2}\ddot{\boldsymbol{p}}\cdot\dot{\boldsymbol{p}}\right) + \frac{1}{2}(\mathrm{e}^t H(t))^{\cdot} = -\rho|\ddot{\boldsymbol{p}}|^2,\end{aligned} \tag{8}$$

so that $E \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}_+)$ and nonincreasing. The function E is defined in such a way that

$$-\frac{\rho}{4}\left(\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2\right)^{\cdot} + \frac{1}{2}H(t) = \mathrm{e}^{-t}E(t). \tag{9}$$

Let us now integrate the latter on (t,T) getting

$$\begin{aligned}&\frac{\rho}{4}\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2 - \frac{\rho}{4}\mathrm{e}^{-T}|\dot{\boldsymbol{p}}(T)|^2 + \frac{1}{2}\int_t^T H(s)\mathrm{d}s = \int_t^T \mathrm{e}^{-s}E(s)\mathrm{d}s \\ &\le E(t)\int_t^T \mathrm{e}^{-s}\mathrm{d}s = E(t)(\mathrm{e}^{-t}-\mathrm{e}^{-T})\end{aligned} \tag{10}$$

where the inequality follows from the monotonicity of E. Hence, by letting $T \to \infty$ in (10) and recalling that $\mathrm{e}^{-T}|\dot{\boldsymbol{p}}(T)|^2 \to 0$, we have proved that

$$\frac{\rho}{4}|\dot{\boldsymbol{p}}(t)|^2 \le E(t) \le E(0). \tag{11}$$

We now turn to the estimate of $E(0)$. At first, note that, by exploiting the bounds (4) and (6) we have that

$$\int_0^1 |\ddot{\boldsymbol{p}}(t)|^2\mathrm{d}t \le \mathrm{e}\int_0^\infty \mathrm{e}^{-t}|\ddot{\boldsymbol{p}}(t)|^2\mathrm{d}t \le \frac{2\mathrm{e}}{\rho}G_\varepsilon[\widehat{\boldsymbol{p}}] \overset{(4)}{\le} \frac{c}{\rho}\varepsilon^2, \tag{12}$$

$$\int_0^1 |\dot{\boldsymbol{p}}(t)|^2\mathrm{d}t \le \mathrm{e}\int_0^\infty \mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2\mathrm{d}t \overset{(6)}{\le} c\varepsilon^2 + \frac{c}{\rho}G_\varepsilon[\widehat{\boldsymbol{p}}] \overset{(4)}{\le} c\left(1+\frac{1}{\rho}\right)\varepsilon^2. \tag{13}$$

In particular, these bounds and $H(t) \le H(0) = G_\varepsilon[\boldsymbol{p}] \le c\varepsilon^2$ suffice in order to conclude that

$$\int_0^1 E(t)\mathrm{d}t \le c(1+\rho)\varepsilon^2. \tag{14}$$

Eventually, by using equality (8) and integrating in time we have

$$
\begin{aligned}
E(0) = \int_0^1 E(0)\mathrm{d}t &\overset{(8)}{=} \int_0^1 \left(E(t)+\rho \int_0^t |\ddot{\boldsymbol{p}}(s)|^2 \mathrm{d}s \right) \mathrm{d}t \\
&\leq \int_0^1 E(t)\mathrm{d}t + \rho \int_0^1 |\ddot{\boldsymbol{p}}(t)|^2 \mathrm{d}t \overset{(14)}{\leq} c(1+\rho)\varepsilon^2. \qquad (15)
\end{aligned}
$$

Going back to (11), we have finally checked the pointwise bound $\rho|\dot{\boldsymbol{p}}(t)|^2 \leq c\varepsilon^2$ and estimate (3) ensues by time rescaling. ☐

2.3 Euler-Lagrange equations

The proof of Theorem 1.1 follows by passing to the limit for $\varepsilon \to 0$ in the Euler-Lagrange equations for the minimizers $\boldsymbol{q}_\varepsilon$ of ER_ε on K_ε. By considering internal variations one has that

$$
0 = \int_0^\infty \left(\rho(\mathrm{e}^{-t/\varepsilon} \ddot{\boldsymbol{q}}_\varepsilon(t))^{\cdot\cdot} + \frac{1}{\varepsilon^2} \mathrm{e}^{-t/\varepsilon} \nabla U(\boldsymbol{q}_\varepsilon(t)) \right) \cdot \boldsymbol{v}(t) \, \mathrm{d}t \qquad (16)
$$

for all $\boldsymbol{v} \in \mathrm{C}_0^\infty(\mathbb{R}_+;\mathbb{R}^m)$. Hence, minimizers of ER_ε solve the Euler-Lagrange equations (2). Indeed, system (2) is solved in the strong sense as the Lipschitz continuity of ∇U and relation (16) entail that

$$
\varepsilon^2 \rho(\mathrm{e}^{-t/\varepsilon} \ddot{\boldsymbol{q}}_\varepsilon(t))^{\cdot\cdot} = -\mathrm{e}^{-t/\varepsilon} \nabla U(\boldsymbol{q}_\varepsilon(t)) \in \mathrm{L}^2(\mathbb{R}_+;\mathbb{R}^m).
$$

In particular, $\boldsymbol{q}_\varepsilon \in \mathrm{H}^4(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$.

2.4 Proof of Theorem 1.1

The pointwise estimate of Lemma 2.2 yields that, by possibly passing to non-relabeled subsequences, we have that $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ locally uniformly. Let us check that $\boldsymbol{q}$ indeed solves the Lagrangian system (1). To this aim, fix any $\boldsymbol{w} \in \mathrm{C}_0^\infty(\mathbb{R}_+;\mathbb{R}^m)$ and choose $\boldsymbol{v}(t) = \boldsymbol{v}_\varepsilon(t) := \mathrm{e}^{t/\varepsilon}\boldsymbol{w}(t)$ in relation (16). As one has that

$$
\ddot{\boldsymbol{v}}_\varepsilon(t) = \mathrm{e}^{t/\varepsilon}\ddot{\boldsymbol{w}}(t) + (2/\varepsilon)\mathrm{e}^{t/\varepsilon}\dot{\boldsymbol{w}}(t) + (1/\varepsilon^2)\mathrm{e}^{t/\varepsilon}\boldsymbol{w}(t),
$$

from (16) we get that

$$
\begin{aligned}
0 &= \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\rho \ddot{\boldsymbol{q}}_\varepsilon(t) \cdot \ddot{\boldsymbol{v}}_\varepsilon(t) + \frac{1}{\varepsilon^2} \nabla U(\boldsymbol{q}_\varepsilon(t)) \cdot \boldsymbol{v}_\varepsilon(t) \right) \mathrm{d}t \\
&= \int_0^\infty \left(\rho \ddot{\boldsymbol{q}}_\varepsilon(t) \cdot \ddot{\boldsymbol{w}}(t) + \frac{2\rho}{\varepsilon} \ddot{\boldsymbol{q}}_\varepsilon(t) \cdot \dot{\boldsymbol{w}}(t) + \frac{\rho}{\varepsilon^2} \ddot{\boldsymbol{q}}_\varepsilon(t) \cdot \boldsymbol{w}(t) + \frac{1}{\varepsilon^2} \nabla U(\boldsymbol{q}_\varepsilon(t)) \cdot \boldsymbol{w}(t) \right) \mathrm{d}t.
\end{aligned}
$$

In particular, one deduces from the latter that

$$
\begin{aligned}
\int_0^\infty \left(\rho \boldsymbol{q}_\varepsilon(t) \cdot \ddot{\boldsymbol{w}}(t) + \nabla U(\boldsymbol{q}_\varepsilon(t)) \cdot \boldsymbol{w}(t) \right) \mathrm{d}t &= \int_0^\infty \left(\varepsilon^2 \rho \dot{\boldsymbol{q}}_\varepsilon(t) \cdot \boldsymbol{w}^{(3)}(t) + 2\varepsilon\rho \dot{\boldsymbol{q}}_\varepsilon(t) \cdot \ddot{\boldsymbol{w}}(t) \right) \mathrm{d}t \\
&= \int_0^T \rho \dot{\boldsymbol{q}}_\varepsilon(t) \cdot \left(\varepsilon^2 \boldsymbol{w}^{(3)}(t) + 2\varepsilon \ddot{\boldsymbol{w}}(t) \right) \mathrm{d}t.
\end{aligned}
$$

By passing to the limit in the latter as $\varepsilon \to 0$ and using the bound (3) we have that

$$\int_0^\infty \big(\rho \boldsymbol{q}(t)\cdot\ddot{\boldsymbol{w}}(t)+\nabla U(\boldsymbol{q}(t))\cdot\boldsymbol{w}(t)\big)\,\mathrm{d}t = 0.$$

Namely, $\boldsymbol{q}$ solves $\rho\ddot{\boldsymbol{q}} = -\nabla U(\boldsymbol{q})$ in the distributional sense. By comparison in the latter we have that $\boldsymbol{q} \in \mathrm{C}^2(\mathbb{R}_+;\mathbb{R}^m)$ so that $\boldsymbol{q}$ is indeed a classical solution of (1). Eventually, as the solution of the second order Cauchy problem (1) is unique (recall that ∇U is Lipschitz continuous), the convergence $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ holds for the whole sequence.

2.5 Integrability conditions at infinity

Before closing this section we shall explicitly remark the crucial role of the two integrability conditions at infinity

$$t \mapsto \mathrm{e}^{-t/\varepsilon}|\boldsymbol{q}|^2,\ t \mapsto \mathrm{e}^{-t/\varepsilon}|\dot{\boldsymbol{q}}|^2 \in \mathrm{L}^1(\mathbb{R}_+;\mathbb{R}^m) \tag{17}$$

which are fulfilled by all trajectories $\boldsymbol{q}$ in K_ε. These conditions correspond to the two *missing* boundary conditions needed in order to complement the fourth-order problem (2). In particular, conditions (17) are responsible for the *non-causality* of the problem at all levels $\varepsilon > 0$: The solution $\boldsymbol{q}$ at time t depends on *future*, i.e., its value on (t,∞). Note however that by taking $\varepsilon \to 0$ causality is restored in the limit, see (1).

In order to illustrate this remark, let us consider once more the scalar linear situation of $U(q) = q^2/2$ and $\rho = 1$. In this case, the solution of $\varepsilon^2 q^{(4)} - 2\varepsilon q^{(3)} + \ddot{q} + q = 0$ can be computed explicitly as $q(t) = \sum_{i=1}^4 c_i \exp(\lambda_{\varepsilon,i}t)$ with

$$\lambda_{\varepsilon,1} = \frac{1-u_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,2} = \frac{1-v_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,3} = \frac{1+u_\varepsilon}{2\varepsilon}, \quad \lambda_{\varepsilon,4} = \frac{1+v_\varepsilon}{2\varepsilon}.$$

In the latter $u_\varepsilon,\ v_\varepsilon \in \mathbb{C}$ are chosen in such a way that $u_\varepsilon^2 = 1-4\varepsilon\mathrm{i}$ and $v_\varepsilon^2 = 1+4\varepsilon\mathrm{i}$, respectively. By exploiting conditions (17) we readily check that, necessarily, $c_3 = c_4 = 0$. Hence, solutions to (2) in fulfilling (17) are of the form $q(t) = c_1\exp(\lambda_{\varepsilon,1}t) + c_2\exp(\lambda_{\varepsilon,2}t)$ and we easily check that $\lambda_{\varepsilon,1} \to i$ and $\lambda_{\varepsilon,2} \to -i$. This corresponds to the fact that the limit of minimizers of ER_ε in K_ε converge to a linear combination of $\sin$ and $\cos$, i.e., a solution of $\ddot{q} + q = 0$.

3 Dissipative evolutions

A distinctive feature of ER variational approach to Lagrangian mechanics resides in its flexibility in encompassing dissipative situations. Indeed, Theorem 1.1 can be quite straightforwardly extended to handle mixed dissipative/nondissipative situations. Let now $\rho \geq 0$ and the *viscosity coefficient* $\nu \geq 0$ be given and consider the functionals

$$\overline{\mathsf{ER}}_\varepsilon[\boldsymbol{q}] := \int_0^\infty \mathrm{e}^{-t/\varepsilon}\left(\frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{q}}(t)|^2 + \frac{\varepsilon\nu}{2}|\dot{\boldsymbol{q}}(t)|^2 + U(\boldsymbol{q}(t))\right)\mathrm{d}t \qquad (\varepsilon > 0).$$

Let $\boldsymbol{q}_\varepsilon$ be the minimizer of $\overline{\mathsf{ER}}_\varepsilon$ on the closed and convex set K_ε which now reads

$$K_\varepsilon := \{\rho\dot{\boldsymbol{q}}, \boldsymbol{q} \in \mathrm{H}^1(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m) \ : \ \boldsymbol{q}(0) = \boldsymbol{q}^0, \ \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1\}.$$

Then, we have the following extension of the ER principle to mixed dissipative/nondissipative situations.

Theorem 3.1 (ER principle, dissipative/nondissipative case)**.** *Assume $\rho+\nu > 0$ and let $\boldsymbol{q}_\varepsilon$ minimize $\overline{\mathsf{ER}}_\varepsilon$ on K_ε. Then, $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ locally uniformly, where*

$$\rho\ddot{\boldsymbol{q}} + \nu\dot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \ \textit{ in } \mathbb{R}_+, \quad \boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1.$$

Note that the very same considerations of Subsection 2.1 can be extended to the present case in order to ensure that such minimizers exist.

3.1 A priori estimate

As for the purely nondissipative case of Theorem 1.1, the convergence proof of Theorem 3.1 follows from an a priori estimate.

Lemma 3.2 (A priori estimate, dissipative/nondissipative case)**.** *Let $\boldsymbol{q}_\varepsilon$ minimize $\overline{\mathsf{ER}}_\varepsilon$ on K_ε. Then,*

$$\rho|\dot{\boldsymbol{q}}_\varepsilon(t)|^2 + \nu\int_0^t |\dot{\boldsymbol{q}}_\varepsilon(s)|^2\,\mathrm{d}s \leq c \quad \forall t > 0. \tag{18}$$

Before proceeding to the proof, let us remark that the two terms in estimate (18) are exactly the ones which are expected in the limit $\varepsilon = 0$. As such, the estimate shows a remarkable optimality with respect to possibly mixed dissipative/nondissipative dynamics. The proof of estimate (18) results by extending the one of Lemma 2.2. In particular, we extend here the argument from [ST10] in order to incorporate dissipative effects.

Proof. We shall reconsider the proof of Lemma 2.2: Letting $\boldsymbol{q}_\varepsilon$ be a minimizer of $\overline{\mathsf{ER}}_\varepsilon$ on K_ε we redefine the rescaled quantities

$$\boldsymbol{p}(t) := \boldsymbol{q}_\varepsilon(\varepsilon t), \quad G_\varepsilon[\boldsymbol{p}] := \int_0^\infty \mathrm{e}^{-t}\left(\frac{\rho}{2}|\ddot{\boldsymbol{p}}(t)|^2 + \frac{\varepsilon\nu}{2}|\dot{\boldsymbol{p}}(t)|^2 + \varepsilon^2 U(\boldsymbol{p}(t))\right)\mathrm{d}t$$

and, accordingly,

$$H(t) := \int_t^\infty \mathrm{e}^{-s}\left(\frac{\rho}{2}|\ddot{\boldsymbol{p}}(s)|^2 + \frac{\varepsilon\nu}{2}|\dot{\boldsymbol{p}}(s)|^2 + \varepsilon^2 U(\boldsymbol{p}(s))\right)\mathrm{d}s.$$

By choosing again $\widehat{\boldsymbol{p}}(t) := \boldsymbol{q}^0 + (\varepsilon\boldsymbol{q}^1)t$ we have that

$$\begin{aligned} G_\varepsilon[\boldsymbol{p}] \leq G_\varepsilon[\widehat{\boldsymbol{p}}] &= \int_0^\infty \mathrm{e}^{-t}\left(\frac{\varepsilon^3\nu}{2}|\boldsymbol{q}^1|^2 + \varepsilon^2 U(\widehat{\boldsymbol{p}}(t))\right)\mathrm{d}t \\ &\leq c\varepsilon^3 + c\varepsilon^2\int_0^\infty \mathrm{e}^{-t}\left(1 + \varepsilon|\boldsymbol{q}^1|t\right)\mathrm{d}t \leq c\varepsilon^2. \end{aligned}$$

In particular, the bound (6) reads in this case as

$$(\rho{+}\varepsilon\nu)\int_0^\infty \mathrm{e}^{-s}|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s \le c\varepsilon^2 + cG_\varepsilon[\boldsymbol{p}] \le c\varepsilon^2. \tag{19}$$

On the other hand, relation (7) in this dissipative/nondissipative context reads

$$\Big(\frac{\rho}{2}\ddot{\boldsymbol{p}}{\cdot}\dot{\boldsymbol{p}}\Big)^{\boldsymbol{\cdot}} = \frac{1}{2}\left(\mathrm{e}^t H(t)\right)^{\boldsymbol{\cdot}} + \rho|\ddot{\boldsymbol{p}}|^2 + \frac{\rho}{2}\ddot{\boldsymbol{p}}{\cdot}\dot{\boldsymbol{p}} + \varepsilon\nu|\dot{\boldsymbol{p}}|^2. \tag{20}$$

Hence, we can redefine the function E as

$$E(t) := \frac{\rho}{4}|\dot{\boldsymbol{p}}(t)|^2 - \frac{\rho}{2}\ddot{\boldsymbol{p}}{\cdot}\dot{\boldsymbol{p}} + \varepsilon\nu\int_0^t |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s + \frac{1}{2}\mathrm{e}^t H(t) \qquad \forall t\ge 0$$

so that, by taking the time derivative and using relation (20), we again have that

$$\dot{E} = -\rho|\ddot{\boldsymbol{p}}|^2. \tag{21}$$

Moreover, we readily check that (see (9))

$$-\frac{\rho}{4}\big(\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2\big)^{\boldsymbol{\cdot}} + \frac{1}{2}H(t) + \varepsilon\nu\mathrm{e}^{-t}\int_0^t |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s = \mathrm{e}^{-t}E(t).$$

Hence, by integrating on (t,T) and using the fact that E is nonincreasing one concludes

$$\begin{aligned}&\frac{\rho}{4}\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2 - \frac{\rho}{4}\mathrm{e}^{-T}|\dot{\boldsymbol{p}}(T)|^2 + \frac{1}{2}\int_t^T H(s)\mathrm{d}s + \varepsilon\nu\int_t^T \mathrm{e}^{-s}\left(\int_0^s |\dot{\boldsymbol{p}}(r)|^2\mathrm{d}r\right)\mathrm{d}s\\ &= \int_t^T \mathrm{e}^{-s}E(s)\mathrm{d}s \le (\mathrm{e}^{-t}{-}\mathrm{e}^{-T})E(t) \le (\mathrm{e}^{-t}{-}\mathrm{e}^{-T})E(0).\end{aligned} \tag{22}$$

Let us now take the limit for $T\to\infty$. By recalling that $\mathrm{e}^{-T}|\dot{\boldsymbol{p}}(T)|^2\to 0$ we get

$$\frac{\rho}{4}\mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2 + \varepsilon\nu\int_t^\infty \mathrm{e}^{-s}\left(\int_0^s |\dot{\boldsymbol{p}}(r)|^2\mathrm{d}r\right)\mathrm{d}s \le \mathrm{e}^{-t}E(0)$$

In particular, $t\mapsto \mathrm{e}^{-t}\int_0^t|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s \in \mathrm{L}^1(\mathbb{R}_+)$ and, owing also to bound (19), it is a standard matter to compute

$$\left(\mathrm{e}^{-t}\int_0^t |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s\right)^{\boldsymbol{\cdot}} = -\mathrm{e}^{-t}\int_0^t |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s + \mathrm{e}^{-t}|\dot{\boldsymbol{p}}(t)|^2$$

and deduce that indeed $t\mapsto \mathrm{e}^{-t}\int_0^t|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s \in \mathrm{W}^{1,1}(\mathbb{R}_+)$. Hence, we also have that $\mathrm{e}^{-t}\int_0^t|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s \to 0$ as $t\to\infty$.

We shall now go back to relation (22), handle the $\varepsilon\nu$-term by

$$\begin{aligned}\varepsilon\nu\int_t^T \mathrm{e}^{-s}\left(\int_0^s |\dot{\boldsymbol{p}}(r)|^2\mathrm{d}r\right)\mathrm{d}s &= -\varepsilon\nu\mathrm{e}^{-T}\int_0^T |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s + \varepsilon\nu\mathrm{e}^{-t}\int_0^t |\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s\\ &\quad + \varepsilon\nu\int_t^T \mathrm{e}^{-s}|\dot{\boldsymbol{p}}(s)|^2\mathrm{d}s,\end{aligned}$$

and take the limit $T \to \infty$ in order to get

$$\frac{\rho}{4}|\dot{\boldsymbol{p}}(t)|^2 + \varepsilon\nu \int_0^t |\dot{\boldsymbol{p}}(s)|^2 \mathrm{d}s \le E(0).$$

By arguing exactly as in (15) we check that $E(0) \le c\varepsilon^2$. Eventually, estimate (18) follows by time rescaling. □

3.2 Proof of Theorem 3.1

We aim now at passing to the limit in the Euler-Lagrange equations

$$0 = \int_0^\infty \left(\varepsilon^2\rho\big(\mathrm{e}^{-t/\varepsilon}\ddot{\boldsymbol{q}}_\varepsilon(t)\big)^{\cdot\cdot} - \varepsilon\nu\big(\mathrm{e}^{-t/\varepsilon}\dot{\boldsymbol{q}}_\varepsilon(t)\big)^{\cdot} + \mathrm{e}^{-t/\varepsilon}\nabla U(\boldsymbol{q}_\varepsilon(t))\right)\cdot \boldsymbol{v}(t)\,\mathrm{d}t \tag{23}$$

for all $\boldsymbol{v} \in \mathrm{C}_0^\infty(\mathbb{R}_+;\mathbb{R}^m)$. By compactness we get that $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ locally uniformly (a posteriori no extraction of subsequence is actually needed here). Fix any $\boldsymbol{w} \in \mathrm{C}_0^\infty(\mathbb{R}_+;\mathbb{R}^m)$ and choose $\boldsymbol{v}(t) = \boldsymbol{v}_\varepsilon(t) := \mathrm{e}^{t/\varepsilon}\boldsymbol{w}(t)$ in relation (23) getting

$$\begin{aligned}
0 &= \int_0^\infty \mathrm{e}^{-t/\varepsilon}\left(\varepsilon^2\rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot\ddot{\boldsymbol{v}}_\varepsilon(t) + \varepsilon\nu\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\dot{\boldsymbol{v}}_\varepsilon(t) + \nabla U(\boldsymbol{q}_\varepsilon(t))\cdot\boldsymbol{v}_\varepsilon(t)\right)\mathrm{d}t\\
&= \int_0^\infty \left(\varepsilon^2\rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot\ddot{\boldsymbol{w}}(t) + 2\varepsilon\rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot\dot{\boldsymbol{w}}(t) + \rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot\boldsymbol{w}(t)\right)\mathrm{d}t\\
&+ \int_0^\infty (\varepsilon\nu\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\dot{\boldsymbol{w}}(t) + \nu\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\boldsymbol{w}(t))\,\mathrm{d}t + \int_0^\infty \nabla U(\boldsymbol{q}_\varepsilon(t))\cdot\boldsymbol{w}(t)\mathrm{d}t.
\end{aligned}$$

Hence, we have proved that

$$\begin{aligned}
&\int_0^\infty \left(\rho\ddot{\boldsymbol{q}}_\varepsilon(t) + \nu\dot{\boldsymbol{q}}_\varepsilon(t) + \nabla U(\boldsymbol{q}_\varepsilon(t))\right)\cdot\boldsymbol{w}(t)\mathrm{d}t\\
&= \int_0^\infty \left(\varepsilon^2\rho\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\boldsymbol{w}^{(3)}(t) + 2\varepsilon\rho\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\ddot{\boldsymbol{w}}(t) - \varepsilon\nu\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\dot{\boldsymbol{w}}(t)\right)\mathrm{d}t\\
&= \int_0^T \rho\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\left(\varepsilon^2\boldsymbol{w}^{(3)}(t) + 2\varepsilon\ddot{\boldsymbol{w}}(t)\right)\mathrm{d}t - \int_0^T \nu\dot{\boldsymbol{q}}_\varepsilon(t)\cdot\varepsilon\dot{\boldsymbol{w}}(t)\,\mathrm{d}t.
\end{aligned}$$

Eventually, by using (18) and by passing to the $\limsup$ as $\varepsilon \to 0$ we have that $\boldsymbol{q}$ solves

$$\rho\ddot{\boldsymbol{q}} + \nu\dot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \ \text{ in } \ \mathbb{R}_+.$$

The check of the initial conditions $\boldsymbol{q}(0) = \boldsymbol{q}^0$ and $\rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1$ is immediate. By uniqueness, the whole sequence $\boldsymbol{q}_\varepsilon$ converges.

3.3 Gradient flows

As a corollary of Theorem 3.1 we have checked the $\varepsilon \to 0$ limit also in the *fully dissipative* situation of *gradient flows*, namely $\rho = 0$ and $\nu > 0$. For the sake of definiteness, we shall record this fact in the following.

Corollary 3.3 (ER principle, gradient flows)**.** *Let $\boldsymbol{q}_\varepsilon$ minimize the functional*

$$\boldsymbol{q} \mapsto \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon\nu}{2}|\dot{\boldsymbol{q}}|^2 + U(\boldsymbol{q}(t))\right) \mathrm{d}t$$

among all trajectories $t \mapsto \boldsymbol{q}(t) \in \mathrm{H}^1(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$ such that $\boldsymbol{q}(0) = \boldsymbol{q}^0$. Then, $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ locally uniformly where $\boldsymbol{q}$ is the unique classical solution of the gradient flow problem

$$\nu\dot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \mathbf{0} \ \textit{ in } \mathbb{R}_+, \quad \boldsymbol{q}(0) = \boldsymbol{q}^0.$$

We shall mention that the limit $\varepsilon \to 0$ in the case of gradient flows has been already tackled by a fairly different approach in [RSSS11b, RSSS11a]. Indeed, in the latter the case of (geodesically) convex [RSSS11b] and semi-convex [RSSS11a] potentials in metric spaces is discussed by a Pontryagin-type argument. In particular, minimizers of the corresponding *metric* version of the functional are proved to converge, up to subsequences, to so-called *curves of maximal slope*.

4 Elliptic regularization on $(0,T)$

Let us now move to the consideration of the finite-time horizon situation. In particular, we shall substitute in time integral on $(0,\infty)$ in the definition of ER_ε (and $\overline{\mathsf{ER}}_\varepsilon$, later) by an integration on $(0,T)$ for some fixed reference time $T > 0$. Namely, we consider the functionals

$$\mathsf{ER}_\varepsilon^T[\boldsymbol{q}] := \int_0^T \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{q}}(t)|^2 + U(\boldsymbol{q}(t))\right) \mathrm{d}t \qquad (\varepsilon > 0)$$

to be minimized on the convex and closed set

$$K^T := \{\rho\dot{\boldsymbol{q}},\ \boldsymbol{q} \in \mathrm{H}^1(0,T;\mathbb{R}^m)\ :\ \boldsymbol{q}(0) = \boldsymbol{q}^0,\ \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1\}.$$

This change brings the ER approach closer to the traditional formulation of the Hamilton principle where some suitable final time is prescribed. The aim of this section is that of reproducing, and in place sharpen, the convergence results of the infinite-time horizon frame of Section 2. Indeed, also in the finite-horizon case $T < \infty$ the limit as $\varepsilon \to 0$ of minimizers of the $\mathsf{ER}_\varepsilon^T$ functional converge to solutions of the Lagrangian system (Theorem 4.2). Moreover, an explicit convergence rate can be exhibited (Theorem 4.3). The latter quantitative error bound is presently not available in the infinite-horizon case.

Note that the convergence proof for minimizers of $\mathsf{ER}_\varepsilon^T$ is substantially different from the corresponding one of the infinite-horizon case. In fact the arguments of Section 2 heavily rely on the invariance of the time-integration interval $\mathbb{R}_+$ with respect to linear time rescalings. Additionally, the appearance of the finiteness of the time interval of integration entails the arising of two final boundary conditions at time T (see (27) below). These final boundary conditions are clearly bound to disappear in the limit $\varepsilon \to 0$. Still, they require specific attention for all $\varepsilon > 0$, exactly in the spirit of Subsection 2.5.

4.1 Well-posedness of the minimum problem

Let us start by checking that indeed minimizers of $\mathsf{ER}_\varepsilon^T$ on K^T exist. In the present finite-time situation the result is even stronger with respect to Lemma 2.1 as the functionals $\mathsf{ER}_\varepsilon^T$ turn out to be uniformly convex for small ε. In particular, the minimum problem is well-posed and minimizers are unique.

Lemma 4.1 (Direct method, $T < \infty$)**.** *Letting ε be small enough, the functional $\mathsf{ER}_\varepsilon^T$ is uniformly convex in $\mathrm{H}^2(0,T;\mathbb{R}^m)$. In particular, $\mathsf{ER}_\varepsilon^T$ admits a unique minimizer in K^T.*

Proof. Recall that $U \in \mathrm{C}^{1,1}$ implies that there exists $\lambda > 0$ such that $\boldsymbol{p}\cdot\mathrm{D}^2U(\boldsymbol{q})\boldsymbol{p} \geq -\lambda|\boldsymbol{p}|^2/2$ for all $\boldsymbol{q},\boldsymbol{p} \in \mathbb{R}^m$. Given $\boldsymbol{q} \in K^T$, consider the function $\boldsymbol{p}(t) := \mathrm{e}^{-t/(2\varepsilon)}\boldsymbol{q}(t)$. We rewrite $\mathsf{ER}_\varepsilon^T[\boldsymbol{q}]$ via $\boldsymbol{p}$ as

$$
\begin{aligned}
\mathsf{ER}_\varepsilon^T[\boldsymbol{q}] &= \int_0^T \left(\frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{p}}(t)|^2+\frac{\rho}{2}|\dot{\boldsymbol{p}}(t)|^2+\frac{\rho-16\varepsilon^2\lambda}{32\varepsilon^2}|\boldsymbol{p}(t)|^2\right)\mathrm{d}t \\
&\quad+\int_0^T\left(\varepsilon\rho\ddot{\boldsymbol{p}}(t)\cdot\dot{\boldsymbol{p}}(t)+\frac{\rho}{4}\ddot{\boldsymbol{p}}(t)\cdot\boldsymbol{p}(t)+\frac{\rho}{4\varepsilon}\dot{\boldsymbol{p}}(t)\cdot\boldsymbol{p}(t)+\mathrm{e}^{-t/\varepsilon}\left(U(\boldsymbol{q}(t))+\frac{\lambda}{2}|\boldsymbol{q}(t)|^2\right)\right)\mathrm{d}t \\
&= \int_0^T \left(\frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{p}}(t)|^2+\frac{\rho}{4}|\dot{\boldsymbol{p}}(t)|^2+\frac{\rho-16\varepsilon^2\lambda}{32\varepsilon^2}\rho|\boldsymbol{p}(t)|^2\right)\mathrm{d}t \\
&\quad+\rho\Big(\varepsilon\dot{\boldsymbol{p}}(T)\cdot\dot{\boldsymbol{p}}(T)-\varepsilon\dot{\boldsymbol{p}}(0)\cdot\dot{\boldsymbol{p}}(0)+\tfrac{1}{4}\dot{\boldsymbol{p}}(T)\cdot\boldsymbol{p}(T)-\tfrac{1}{4}\dot{\boldsymbol{p}}(0)\cdot\boldsymbol{p}(0)+\tfrac{1}{2\varepsilon}|\boldsymbol{p}(T)|^2-\tfrac{1}{2\varepsilon}|\boldsymbol{p}(0)|^2\Big) \\
&\quad+\int_0^T\mathrm{e}^{-t/\varepsilon}\left(U(\boldsymbol{q}(t))+\frac{\lambda}{2}|\boldsymbol{q}(t)|^2\right)\mathrm{d}t \\
&=: A_\varepsilon[\boldsymbol{p}]+B_\varepsilon[\boldsymbol{p}]+C_\varepsilon[\boldsymbol{q}].
\end{aligned}
$$

Here, A_ε is quadratic and uniformly convex (of constant $\alpha_\varepsilon > 0$, say) with respect to $\boldsymbol{p}$ in $\mathrm{H}^2(0,T;\mathbb{R}^m)$ for all $\varepsilon < (16\lambda)^{-1/2}$ and C_ε is clearly convex with respect to $\boldsymbol{q}$. The same holds also for the functional B_ε for an elementary computation ensures that

$$
B_\varepsilon[\boldsymbol{p}] = \frac{3\varepsilon\rho}{8}\mathrm{e}^{-T/\varepsilon}|\dot{\boldsymbol{q}}(T)|^2+\frac{\varepsilon\rho}{8}\mathrm{e}^{-T/\varepsilon}\,|\dot{\boldsymbol{q}}(T)-\boldsymbol{q}(T)/\varepsilon|^2-\frac{3\varepsilon\rho}{8}|\boldsymbol{q}^1|^2-\frac{\varepsilon\rho}{8}|\boldsymbol{q}^1-\boldsymbol{q}^0/\varepsilon|^2.
$$

Let now $\theta \in [0,1]$, $\boldsymbol{q}_0,\boldsymbol{q}_1 \in K^T$, and define accordingly $\boldsymbol{p}_0,\boldsymbol{p}_1$ as above. We have that

$$
\begin{aligned}
\mathsf{ER}_\varepsilon^T[(1-\theta)\boldsymbol{q}_0+\theta\boldsymbol{q}_1] &= A_\varepsilon[(1-\theta)\boldsymbol{p}_0+\theta\boldsymbol{p}_1]+B_\varepsilon[(1-\theta)\boldsymbol{p}_0+\theta\boldsymbol{p}_1]+C_\varepsilon[(1-\theta)\boldsymbol{q}_0+\theta\boldsymbol{q}_1] \\
&\leq -\frac{\alpha_\varepsilon}{2}\theta(1-\theta)\|\boldsymbol{p}_0-\boldsymbol{p}_1\|_{\mathrm{H}^2}^2+(1-\theta)\mathsf{ER}_\varepsilon^T[\boldsymbol{q}_0]+\theta\mathsf{ER}_\varepsilon^T[\boldsymbol{q}_1]
\end{aligned}
$$

and the assertion follows as $\|\boldsymbol{p}_0-\boldsymbol{p}_1\|_{\mathrm{H}^2}^2 \geq \varepsilon^4\mathrm{e}^{-T/\varepsilon}\|\boldsymbol{q}_0-\boldsymbol{q}_1\|_{\mathrm{H}^2}^2$. □

4.2 Convergence of minimizers

The main result of this section is the following.

Theorem 4.2 (ER principle, $T < \infty$)**.** *Let $\boldsymbol{q}_\varepsilon$ minimize* $\mathrm{ER}_\varepsilon^T$ *in K^T. Then, $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ uniformly where $\boldsymbol{q}$ solves the Lagrangian system*

$$\rho\ddot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \text{ in } (0,T), \quad \boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1. \tag{24}$$

Before moving on to the proof of Theorem 4.2, let us specify the Euler-Lagrange equations for the minimizers $\boldsymbol{q}_\varepsilon$ of $\mathrm{ER}_\varepsilon^T$ on K^T. In particular, one has that

$$\begin{aligned} 0 &= \rho \mathrm{e}^{-T/\varepsilon}\ddot{\boldsymbol{q}}_\varepsilon(T)\cdot\dot{\boldsymbol{v}}(T) - \rho(\mathrm{e}^{-t/\varepsilon}\ddot{\boldsymbol{q}}_\varepsilon)^{\cdot}(T)\cdot\boldsymbol{v}(T) \\ &\quad + \int_0^T \left(\rho(\mathrm{e}^{-t/\varepsilon}\ddot{\boldsymbol{q}}_\varepsilon(t))^{\cdot\cdot} + \frac{1}{\varepsilon^2}\mathrm{e}^{-t/\varepsilon}\nabla U(\boldsymbol{q}_\varepsilon(t))\right)\cdot\boldsymbol{v}(t)\,\mathrm{d}t \end{aligned}$$

for all $\boldsymbol{v} \in \mathrm{C}_0^\infty(0,T;\mathbb{R}^m)$ and hence

$$\varepsilon^2\rho\boldsymbol{q}^{(4)} - 2\varepsilon\rho\boldsymbol{q}^{(3)} + \rho\ddot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \text{ in } (0,T), \tag{25}$$

$$\boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1, \tag{26}$$

$$\rho\ddot{\boldsymbol{q}}(T) = \rho\boldsymbol{q}^{(3)}(T) = \boldsymbol{0}. \tag{27}$$

Note the occurrence of the two extra final boundary conditions (27) at time T. These conditions will disappear in the limit $\varepsilon \to 0$, see (24).

Proof. One has to start by establishing uniform estimates on $\boldsymbol{q}_\varepsilon$ in the spirit of Lemma 2.2, although necessarily by a different technique. We follow here the argument of [Ste11] and perform some modifications in order to cope with the possible nonconvexity of U (the original argument from [Ste11] works for convex potentials only). Take the scalar product of equation (25) and $\dot{\boldsymbol{q}}_\varepsilon - \boldsymbol{q}^1$ and integrate on $(0,t)$ getting

$$\begin{aligned} 0 &= \varepsilon^2\rho\boldsymbol{q}_\varepsilon^{(3)}(t)\cdot(\dot{\boldsymbol{q}}_\varepsilon(t) - \boldsymbol{q}^1) - \frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{q}}_\varepsilon(t)|^2 + \frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{q}}_\varepsilon(0)|^2 - 2\varepsilon\rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot(\dot{\boldsymbol{q}}_\varepsilon(t) - \boldsymbol{q}^1) \\ &\quad + 2\varepsilon\rho\int_0^t |\ddot{\boldsymbol{q}}_\varepsilon(s)|^2\,\mathrm{d}s + \frac{\rho}{2}|\dot{\boldsymbol{q}}_\varepsilon(t) - \boldsymbol{q}^1|^2 + U(\boldsymbol{q}_\varepsilon(t)) - U(\boldsymbol{q}^0) + \int_0^t \nabla U(\boldsymbol{q}_\varepsilon(s))\cdot\boldsymbol{q}^1\mathrm{d}s. \end{aligned} \tag{28}$$

Now, we integrate (28) on $(0,T)$ and use the final boundary conditions (27) in order to get that

$$\begin{aligned} 0 &= -\frac{3\varepsilon^2\rho}{2}\int_0^T |\ddot{\boldsymbol{q}}_\varepsilon(t)|^2\,\mathrm{d}t + \frac{\varepsilon^2 T\rho}{2}|\ddot{\boldsymbol{q}}_\varepsilon(0)|^2 - \varepsilon\rho|\dot{\boldsymbol{q}}_\varepsilon(T) - \boldsymbol{q}^1|^2 + 2\varepsilon\rho\int_0^T\int_0^t |\ddot{\boldsymbol{q}}_\varepsilon(s)|^2\,\mathrm{d}s\,\mathrm{d}t \\ &\quad + \frac{\rho}{2}\int_0^T |\dot{\boldsymbol{q}}_\varepsilon(t) - \boldsymbol{q}^1|^2\mathrm{d}t + \int_0^T U(\boldsymbol{q}_\varepsilon(t))\mathrm{d}t - TU(\boldsymbol{q}^0) + \int_0^T\int_0^t \nabla U(\boldsymbol{q}_\varepsilon(s))\cdot\boldsymbol{q}^1\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{29}$$

Finally, add (29) to (28) with $t = T$ and use again the boundary conditions (27) getting

$$\begin{aligned} &\left(2\varepsilon - \frac{3\varepsilon^2}{2}\right)\int_0^T \rho\,|\ddot{\boldsymbol{q}}_\varepsilon(t)|^2\,\mathrm{d}t + \frac{\varepsilon^2(1+T)}{2}\rho\,|\ddot{\boldsymbol{q}}_\varepsilon(0)|^2 + \left(\frac{1}{2} - \varepsilon\right)\rho|\dot{\boldsymbol{q}}_\varepsilon(T) - \boldsymbol{q}^1|^2 \\ &+ 2\varepsilon\rho\int_0^T\int_0^t |\ddot{\boldsymbol{q}}_\varepsilon(s)|^2\,\mathrm{d}s\,\mathrm{d}t + \frac{\rho}{2}\int_0^T |\dot{\boldsymbol{q}}_\varepsilon(t) - \boldsymbol{q}^1|^2\mathrm{d}t + U(\boldsymbol{q}_\varepsilon(T)) + \int_0^T U(\boldsymbol{q}_\varepsilon(t))\mathrm{d}t \\ &\le c(T) + \int_0^T \nabla U(\boldsymbol{q}_\varepsilon(t))\cdot\boldsymbol{q}^1\mathrm{d}t + \int_0^T\int_0^t \nabla U(\boldsymbol{q}_\varepsilon(s))\cdot\boldsymbol{q}^1\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{30}$$

By recalling that $U \in \mathrm{C}^{1,1}$ it is a standard matter to bound the latter right-hand side by $c(T) + (1/4)\|\dot{\boldsymbol{q}}_\varepsilon - \boldsymbol{q}^1\|^2_{\mathrm{L}^2}$ so that we have

$$\rho\|\dot{\boldsymbol{q}}_\varepsilon\|^2_{\mathrm{H}^1} \le c(T). \tag{31}$$

Hence, by possibly passing to non-relabeled subsequences, we have that $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ uniformly. Eventually, we check that $\boldsymbol{q}$ indeed classically solves the Lagrangian system (25) by arguing along the lines of Subsection 3.2. ☐

4.3 Quantitative error bound.

As already mentioned, in the finite-time case $T < \infty$ the convergence result of Theorem 4.2 can be refined in order to yield a quantitative rate estimate.

Theorem 4.3 (Error control, $T < \infty$)**.** *Under the assumptions of Theorem* 4.2 *we have that* $\rho\|\boldsymbol{q} - \boldsymbol{q}_\varepsilon\|_{\mathrm{H}^{1+\eta}} \le c(T)\varepsilon^{(1-\eta)/2}$ *for all* $\eta \in [0,1)$.

Proof. The argument relies on establishing an extra estimate. From bound (31) and the Lipschitz continuity of ∇U we have that $\varepsilon^2\rho\boldsymbol{q}_\varepsilon^{(4)} - 2\varepsilon\rho\boldsymbol{q}_\varepsilon^{(3)} + \rho\ddot{\boldsymbol{q}}_\varepsilon$ is uniformly bounded in $\mathrm{L}^2(0,T;\mathbb{R}^m)$, depending on T. Hence, by integrating its squared norm we have that

$$\begin{aligned}
&\varepsilon^4\int_0^T \rho|\boldsymbol{q}_\varepsilon^{(4)}(t)|^2\mathrm{d}t + 4\varepsilon^2\int_0^T \rho|\boldsymbol{q}_\varepsilon^{(3)}(t)|^2\mathrm{d}t + \int_0^T \rho|\ddot{\boldsymbol{q}}_\varepsilon(t)|^2\mathrm{d}t\\
&\le c(T) + 2\varepsilon^3\int_0^T \rho\boldsymbol{q}_\varepsilon^{(4)}(t)\cdot\boldsymbol{q}_\varepsilon^{(3)}(t)\mathrm{d}t + 2\varepsilon\int_0^T \rho\boldsymbol{q}_\varepsilon^{(3)}(t)\cdot\ddot{\boldsymbol{q}}_\varepsilon(t)\mathrm{d}t - \varepsilon^2\int_0^T \rho\boldsymbol{q}_\varepsilon^{(4)}(t)\cdot\ddot{\boldsymbol{q}}_\varepsilon(t)\mathrm{d}t\\
&\overset{(27)}{=} c(T) - \varepsilon^3\rho|\boldsymbol{q}_\varepsilon^{(3)}(0)|^2 - \varepsilon\rho|\ddot{\boldsymbol{q}}_\varepsilon(0)|^2 + \varepsilon^2\rho\boldsymbol{q}_\varepsilon^{(3)}(0)\cdot\ddot{\boldsymbol{q}}_\varepsilon(0) + 2\varepsilon^2\int_0^T \rho|\boldsymbol{q}_\varepsilon^{(3)}(t)|^2\mathrm{d}t.
\end{aligned}$$

This entails that $\varepsilon^2\rho^{1/2}\boldsymbol{q}_\varepsilon^{(4)}$, $\varepsilon\rho^{1/2}\boldsymbol{q}_\varepsilon^{(3)}$, and $\rho^{1/2}\ddot{\boldsymbol{q}}_\varepsilon$ are bounded in $\mathrm{L}^2(0,T;\mathbb{R}^m)$. Moreover, the Gagliardo-Nirenberg inequality ensures that

$$\begin{aligned}
&\rho^{1/2}\|\boldsymbol{q}_\varepsilon^{(3)}\|_{\mathrm{L}^\infty} \le c(T)\big(\rho^{1/2}\|\boldsymbol{q}_\varepsilon^{(3)}\|_{\mathrm{L}^2} + \rho^{1/2}\|\boldsymbol{q}_\varepsilon^{(3)}\|^{1/2}_{\mathrm{L}^2}\|\boldsymbol{q}_\varepsilon^{(4)}\|^{1/2}_{\mathrm{L}^2}\big) \le c(T)\left(\frac{1}{\varepsilon} + \frac{1}{\varepsilon^{3/2}}\right),\\
&\rho^{1/2}\|\ddot{\boldsymbol{q}}_\varepsilon\|_{\mathrm{L}^\infty} \le c(T)\left(1 + \frac{1}{\varepsilon}\right).
\end{aligned} \tag{32}$$

Take now the difference between (24) and (25), test it on $\dot{\boldsymbol{p}}_\varepsilon := \dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}_\varepsilon$, and integrate on $(0,t)$ getting

$$\begin{aligned}
&\frac{\rho}{2}|\dot{\boldsymbol{p}}_\varepsilon(t)|^2 = -\varepsilon^2\int_0^t \rho\boldsymbol{q}_\varepsilon^{(4)}(s)\cdot\dot{\boldsymbol{p}}_\varepsilon(s)\mathrm{d}s + 2\varepsilon\int_0^t \rho\boldsymbol{q}_\varepsilon^{(3)}(s)\cdot\dot{\boldsymbol{p}}_\varepsilon(s)\mathrm{d}s\\
&- \int_0^t \big(\nabla U(\boldsymbol{q}(s)) - \nabla U(\boldsymbol{q}_\varepsilon(s))\big)\cdot\dot{\boldsymbol{p}}_\varepsilon(s)\mathrm{d}s\\
&\le -\varepsilon^2\rho\boldsymbol{q}_\varepsilon^{(3)}(t)\cdot\dot{\boldsymbol{p}}_\varepsilon(t) + \varepsilon^2\int_0^t \rho\boldsymbol{q}_\varepsilon^{(3)}(s)\cdot\ddot{\boldsymbol{p}}_\varepsilon(s)\mathrm{d}s + 2\varepsilon\rho\ddot{\boldsymbol{q}}_\varepsilon(t)\cdot\dot{\boldsymbol{p}}_\varepsilon(t) - 2\varepsilon\int_0^t \rho\ddot{\boldsymbol{q}}_\varepsilon(s)\cdot\ddot{\boldsymbol{p}}_\varepsilon(s)\mathrm{d}s\\
&+ c\int_0^t \rho|\boldsymbol{p}_\varepsilon(s)||\dot{\boldsymbol{p}}_\varepsilon(s)|\mathrm{d}s \overset{(32)}{\le} c(T)\varepsilon + \frac{\rho}{4}|\dot{\boldsymbol{p}}_\varepsilon(t)|^2 + c(T)\int_0^t \rho|\dot{\boldsymbol{p}}_\varepsilon(s)|^2\mathrm{d}s
\end{aligned}$$

so that by means of the Gronwall Lemma we get that $\rho\|\dot{\boldsymbol{q}}-\dot{\boldsymbol{q}}_\varepsilon\|^2_{\mathrm{L}^\infty} \le c(T)\varepsilon$. By interpolation [BL76], for all $\eta \in (0,1)$ we have $\rho\|\boldsymbol{q}-\boldsymbol{q}_\varepsilon\|_{(\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,1}} \le c(T)\|\dot{\boldsymbol{q}}-\dot{\boldsymbol{q}}_\varepsilon\|^{1-\eta}_{\mathrm{L}^\infty}\|\boldsymbol{q}-\boldsymbol{q}_\varepsilon\|^{\eta}_{\mathrm{H}^2} \le c(T)\varepsilon^{(1-\eta)/2}$ (which is stronger than the statement). Eventually, we conclude by noting that $(\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,1} \subset (\mathrm{W}^{1,\infty},\mathrm{H}^2)_{\eta,2} \subset (\mathrm{H}^1,\mathrm{H}^2)_{\eta,2} = \mathrm{H}^{1+\eta}$ with continuous injections. ☐

4.4 More general potentials.

Before closing this section let us mention that the above results are valid also for some more general choices of the potential U.

First of all, the regularity requirements on the potential U could be slightly weakened. Given the sequence $\boldsymbol{q}_{\varepsilon_n}$ of minimizers of $\mathsf{ER}^T_{\varepsilon_n}$ on K^T as $\varepsilon_n \to 0$, the convergence result of Theorem 4.2 is valid, up to the possible extraction of subsequences, by relaxing the requirement $U \in \mathrm{C}^{1,1}(\mathbb{R}^m)$ for instance as $U \in \mathrm{C}^1(\mathbb{R}^m)$ and $|\nabla U(\boldsymbol{q})| \le c(1+|\boldsymbol{q}|^2)$ for all $\boldsymbol{q} \in \mathbb{R}^m$.

Moreover, the case of a time-dependent potential $U : \mathbb{R}^m \times [0,T] \to \mathbb{R}$ can also be accommodated within the theory along as the dependence on time is sufficiently smooth. By assuming for instance that both $\nabla_{\boldsymbol{q}} U$ and $\partial_t U$ are Lipschitz continuous with respect to $\boldsymbol{q}$ uniformly in t, it is a standard matter to reproduce the argument of Subsection 4.2 and control the extra term $\int_0^T \partial_t U(\boldsymbol{q}_\varepsilon(s),s)\mathrm{d}s + \int_0^T \int_0^t \partial_t U(\boldsymbol{q}_\varepsilon(s),s)\mathrm{d}s\,\mathrm{d}t$ which would appear in the right-hand side of (30). Moreover, Theorem 4.3 would remain valid with unchanged proof.

4.5 Dissipative evolutions.

Also in the finite-time case, the convergence result of Theorem 4.2 can be extended to mixed dissipative/nondissipative situations. In particular, by letting $\rho,\ \nu \ge 0$ one considers the minimization of the functionals

$$\overline{\mathsf{ER}}^T_\varepsilon[\boldsymbol{q}] := \int_0^T \mathrm{e}^{-t/\varepsilon}\left(\frac{\varepsilon^2\rho}{2}|\ddot{\boldsymbol{q}}(t)|^2 + \frac{\varepsilon\nu}{2}|\dot{\boldsymbol{q}}(t)|^2 + U(\boldsymbol{q}(t))\right)\mathrm{d}t \qquad (\varepsilon > 0)$$

over the convex set K^T. By assuming $\rho+\nu > 0$ and letting ε be small enough the same results of Lemma 4.1 hold. In particular, $\overline{\mathsf{ER}}^T_\varepsilon$ is uniformly convex hence admitting a unique minimizer on K^T. Moreover, we have the following.

Theorem 4.4 (ER principle, dissipative/nondissipative case, $T < \infty$)**.** *Let $\rho+\nu > 0$ and $\boldsymbol{q}_\varepsilon$ minimize $\overline{\mathsf{ER}}^T_\varepsilon$ in K^T. Then, $\boldsymbol{q}_\varepsilon \to \boldsymbol{q}$ uniformly where*

$$\rho\ddot{\boldsymbol{q}} + \nu\dot{\boldsymbol{q}} + \nabla U(\boldsymbol{q}) = \boldsymbol{0} \ \text{ in } (0,T), \quad \boldsymbol{q}(0) = \boldsymbol{q}^0, \quad \rho\dot{\boldsymbol{q}}(0) = \rho\boldsymbol{q}^1.$$

We shall not present here the detailed proof of the latter as it can be obtained along the very same lines (and some additional technicalities) of the proof of Theorem 3.1. Some detail in this direction is however provided in the forthcoming [LS11] where some infinite-dimensional PDE situation is discussed. The conclusions of Theorem 4.3 hold unchanged as long as $\rho > 0$ and the proof is indeed an extension of the proposed one. For $\rho = 0$ one resorts in the (necessarily weaker) quantitative convergence result $\nu\|\boldsymbol{q}-\boldsymbol{q}_\varepsilon\|_{\mathrm{H}^\eta} \le c(T)\varepsilon^{(1-\eta)/2}$ for $\eta \in [0,1)$.

5 Γ-convergence

The ER variational formalism is well-suited in order to describe limiting behaviors. In particular, starting from the mixed dissipative/nondissipative situation of Section 3, we shall here comment on the possibility of considering from a variational viewpoint the limits $\rho \to 0$ and $\nu \to 0$. This will be done within the classical frame of Γ-convergence [Dal93, DF79]. Additionally, we will prove that, under suitable specifications, the finite-horizon problem Γ-converges to the infinite-horizon problem as $T \to \infty$.

Let us mention that all the Γ-limits are taken for constant ε as combined Γ-convergence analyses for both parameters and $\varepsilon \to 0$ are currently not available. The reader is however referred to MIELKE & ORTIZ [MO08] and [AS11, MS08] for some Γ-convergence result on ER-type functionals in the doubly nonlinear parabolic setting.

5.1 Viscous Γ-limit $\rho \to 0$

We start by defining the functionals F^ρ over $\mathrm{H}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$ for $\rho \geq 0,\ \nu > 0$ as

$$F^\rho[\boldsymbol{q}] = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2 \rho}{2} |\ddot{\boldsymbol{q}}(t)|^2 + \frac{\varepsilon \nu}{2} |\dot{\boldsymbol{q}}(t)|^2 + U(\boldsymbol{q}(t)) \right) \mathrm{d}t \text{ if } \boldsymbol{q} \in K_\varepsilon^\rho \text{ and } F^\rho[\boldsymbol{q}] = \infty \text{ else}$$

where the notation $K_\varepsilon^\rho := K_\varepsilon$ is just intended to stress that indeed the constraint depends on ρ as well. Our result reads as follows.

Lemma 5.1 (Γ-limit $\rho \to 0$). *We have that $F^\rho \xrightarrow{\Gamma} F^0$ weakly in $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$.*

Proof. Given $\boldsymbol{q} \in K_\varepsilon^0$ one can use singular perturbations in order to find a sequence $\boldsymbol{q}^\rho \in K_\varepsilon^\rho$ with $\boldsymbol{q}^\rho \to \boldsymbol{q}$ strongly in K_ε^0 such that $\rho \int_0^\infty \mathrm{e}^{-t/\varepsilon} |\ddot{\boldsymbol{q}}^\rho(t)|^2 \mathrm{d}t \to 0$. On the other hand, let $\boldsymbol{q}^\rho \to \boldsymbol{q}$ weakly in $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$. As $F^0 \leq F^\rho$, we readily check that $F^0[\boldsymbol{q}] \leq \liminf_{\rho \to 0} F^0[\boldsymbol{q}^\rho] \leq \liminf_{\rho \to 0} F^\rho[\boldsymbol{q}^\rho]$. □

5.2 Nondissipative Γ-limit $\nu \to 0$

In order to formalize our Γ-convergence result, we introduce the functionals F^ν for $\rho > 0$ and $\nu \geq 0$ as

$$F^\nu[\boldsymbol{q}] = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{\varepsilon^2 \rho}{2} |\ddot{\boldsymbol{q}}(t)|^2 + \frac{\varepsilon \nu}{2} |\dot{\boldsymbol{q}}(t)|^2 + U(\boldsymbol{q}(t)) \right) \mathrm{d}t \text{ if } \boldsymbol{q} \in K_\varepsilon \text{ and } F^\nu[\boldsymbol{q}] = \infty \text{ else.}$$

We have the following.

Lemma 5.2 (Γ-limit $\nu \to 0$). *We have that $F^\nu \xrightarrow{\Gamma} F^0$ weakly in $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$.*

Proof. The existence of a recovery sequence is immediate by pointwise convergence since $F^\nu[\boldsymbol{q}] \to F^0[\boldsymbol{q}]$ for all $\boldsymbol{q} \in K_\varepsilon$. Moreover, we have that $F^0 \leq F^\nu$ and we readily check that $F^0[\boldsymbol{q}] \leq \liminf_{\nu \to 0} F^0[\boldsymbol{q}^\nu] \leq \liminf_{\nu \to 0} F^\nu[\boldsymbol{q}^\nu]$ and the assertion follows. □

5.3 Infinite-horizon Γ-limit $T \to \infty$

We shall be considering all functionals ER_ε and $\mathsf{ER}_\varepsilon^T$ to be defined on the common space $\mathrm{H}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$ and specify, for all $t \in (0, \infty]$,

$$F^t[\boldsymbol{q}] := \mathsf{ER}_\varepsilon^t[\boldsymbol{q}] \ \text{ if } \ \boldsymbol{q} \in K_\varepsilon \ \text{ and } \ F^t[\boldsymbol{q}] = \infty \ \text{ else.}$$

Hence, our result reads as follows.

Lemma 5.3 (Γ-limit $T \to \infty$)**.** *We have that* $F^T \xrightarrow{\Gamma} F^\infty$ *weakly in* $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$.

Proof. The existence of a recovery sequence ensues from the pointwise convergence $F^T[\boldsymbol{q}] \to F^\infty[\boldsymbol{q}]$ for all $\boldsymbol{q} \in K_\varepsilon$. Assume now to be given $\boldsymbol{q}^T \to \boldsymbol{q}^\infty$ weakly in $\mathrm{L}^2(\mathbb{R}_+, \mathrm{e}^{-t/\varepsilon}\mathrm{d}t; \mathbb{R}^m)$. By taking with no loss of generality $\liminf_{T\to\infty} F^T[\boldsymbol{q}^T] < \infty$ we have that

$$\liminf_{T\to\infty} \int_0^T \mathrm{e}^{-t/\varepsilon} |\ddot{\boldsymbol{q}}^T(t)|^2 \mathrm{d}t \geq \int_0^\infty \mathrm{e}^{-t/\varepsilon} |\ddot{\boldsymbol{q}}^\infty(t)|^2 \mathrm{d}t$$

and $\boldsymbol{q}^T \to \boldsymbol{q}^\infty$ pointwise almost everywhere. Eventually, $F^T[\boldsymbol{q}^T] \to F^\infty[\boldsymbol{q}^\infty]$ by Dominated Convergence as U is Lipschitz. □

6 Time discretization

We collect in this section some remark on suitable time-discrete versions of the ER principle. Let us focus first on the finite-time case of Section 4. Setting the time step $\tau := T/n \, (n \in \mathbb{N})$, we consider the time-discrete functionals

$$\mathsf{ER}_{\varepsilon\tau}[\boldsymbol{q}_0, \dots, \boldsymbol{q}_n] = \sum_{j=2}^{m} \tau \mathrm{e}_{\varepsilon\tau,j} \frac{\varepsilon^2 \rho}{2} |\delta^2 \boldsymbol{q}_j|^2 + \sum_{j=2}^{n-1} \tau \mathrm{e}_{\varepsilon\tau,j+2} U(\boldsymbol{q}_j).$$

Given $(\boldsymbol{q}_0, \dots, \boldsymbol{q}_n)$, in the latter we have indicated with $\delta\boldsymbol{q}$ its *discrete derivative*, namely $\delta\boldsymbol{q}_j := (\boldsymbol{q}_j - \boldsymbol{q}_{j-1})/\tau$, $\delta^2\boldsymbol{q} = \delta(\delta\boldsymbol{q})$, and so on. The weights $\mathrm{e}_{\varepsilon\tau,j}$ are given by $\mathrm{e}_{\varepsilon\tau,j} := (\varepsilon/(\varepsilon+\tau))^j$ and play the role of the decaying weight $t \mapsto \mathrm{e}^{-t/\varepsilon}$ in the discrete setting. In particular, note that $\mathrm{e}_{\varepsilon\tau,0} = 1$ and $\delta \mathrm{e}_{\varepsilon\tau,j} + \mathrm{e}_{\varepsilon\tau,j}/\varepsilon = 0$. Namely, $\mathrm{e}_{\varepsilon\tau,j}$ is the implicit Euler discretization of the Cauchy problem $\dot{w} + w/\varepsilon = 0$ and $w(0) = 1$.

We shall be concerned with minimizing $\mathsf{ER}_{\varepsilon\tau}$ over the discrete analog of K_ε that is

$$K_{\varepsilon\tau} := \big\{ (\boldsymbol{q}_0, \dots, \boldsymbol{q}_n) \ : \ \boldsymbol{q}_0 = \boldsymbol{q}^0, \ \rho\delta\boldsymbol{q}_1 = \rho\boldsymbol{q}^1 \big\}.$$

It can be proved that, at least for small ε, this minimization problem has a unique solution [LS11]. Moreover, the minimizer fulfills the discrete Euler-Lagrange system

$$\varepsilon^2 \rho \delta^4 \boldsymbol{q}_{j+2} - 2\varepsilon\rho\delta^3 \boldsymbol{q}_{j+1} + \rho\delta^2 \boldsymbol{q}_j + \nabla U(\boldsymbol{q}_j) = 0 \quad j = 2, \dots, n-2, \tag{33}$$

$$\boldsymbol{q}_0 = \boldsymbol{q}^0, \ \rho\delta\boldsymbol{q}_1 = \rho\boldsymbol{q}^1, \tag{34}$$

$$\rho\delta^2 \boldsymbol{q}_n = \rho\delta^3 \boldsymbol{q}_n = \boldsymbol{0}. \tag{35}$$

This scheme is proved to be unconditionally stable and convergent in [Ste11] and can be easily extended in order to cope with the dissipative case of Subsection 3 (see [LS11]).

The system (33)-(35) can be regarded as the *variational integrator* [HLW06] corresponding to the ER principle. We shall stress that the scheme (33)-(35) is computationally more expensive (a system of $n \times d$ nonlinear equations) with respect to the classical implicit Euler scheme (corresponding to $\varepsilon = 0$ in (33), n systems of d nonlinear equations), not speaking of explicit or symplectic Euler (direct substitution) [HLW06]. Indeed, for all $\varepsilon > 0$ the time-discrete ER principle is noncausal and a full system over the time indices has to be solved. This is particularly critical for final conditions (35) are crucially entering the picture. An illustration of the convergence of the scheme is given in Figure 2.

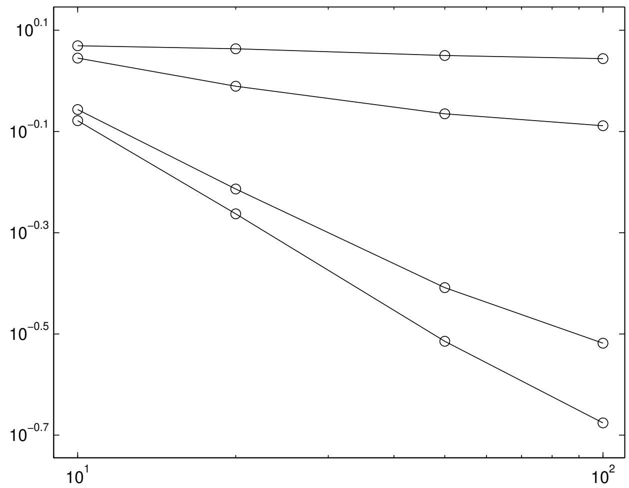

Figure 2: Convergence in the nonlinear case of $U(q) = q^4/4$, $q^0 = 1$, $q^1 = 0$, $T = 10$. The figure plots in log-log scale the error in the uniform norm $\sup_{[0,T]} |\boldsymbol{q} - \boldsymbol{q}_{\varepsilon\tau}|$ against $1/\tau$. The different error curves correspond to the different choices $\varepsilon = 0.2, 0.1, 0.05, 0.02$ (top to bottom).

A remarkable trait of the scheme (33)-(35) is however that of showing some additional stability for $\varepsilon > 0$. In particular, some *explicit* version of the scheme (33)-(35) (i.e., replacing $\nabla U(\boldsymbol{q}_j)$ with $\nabla U(\boldsymbol{q}_{j-1})$ in (33)) shows conditional stability. This contrasts with the instability of the explicit Euler scheme.

Let us mention that the infinite-horizon situation $T = \infty$ seems less amenable from the numerical viewpoint. This is due to the fact that the final conditions (35) have to be replaced with specific summability conditions at infinity as commented in Subsection 2.5. In order to avoid solving an infinite system of equations, one might consider imposing two extra initial conditions such that the above mentioned summability is met in a sort of a *shooting* strategy. As the linear case of Subsection 2.5 show, this turns however turns out to be a tricky task.

Before closing this section let us mention that the same drawback is of course exhibited also by the modifications of (33) given by

$$\varepsilon^2 \rho \delta^4 \boldsymbol{q}_j - 2\varepsilon\rho\delta^3 \boldsymbol{q}_j + \rho\delta^2 \boldsymbol{q}_j + \nabla U(\boldsymbol{q}_i) = 0 \quad \text{for } i = j,\ j-1,\ j-2.$$

Note that the latter schemes cannot be obtained as Euler-Lagrange equations of (variants of) the functionals $\mathrm{ER}_{\varepsilon\tau}$.

Acknowledgements

U.S. and M.L. are partly supported by FP7-IDEAS-ERC-StG Grant # 200947 *BioSMA*. U.S. acknowledges the partial support of CNR-AVČR Grant *SmartMath*, and the Alexander von Humboldt Foundation. Furthermore, M.L. thanks the IMATI Pavia, where part of the work was conducted, for its kind hospitality.

References

[Arn89] V. I. Arnol′d. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1989. Translated from the Russian by K. Vogtmann and A. Weinstein.

[AS10] G. Akagi and U. Stefanelli. A variational principle for doubly nonlinear evolution. *Appl. Math. Lett.*, 23(9):1120–1124, 2010.

[AS11] G. Akagi and U. Stefanelli. Weighted energy-dissipation functionals for doubly nonlinear evolution. *J. Funct, Anal.*, 260(9):2541–2578, 2011.

[Bas07] J.-L. Basdevant. *Variational principles in physics*. Springer, New York, 2007.

[Ber09] V. L. Berdichevsky. *Variational principles of continuum mechanics. I*. Interaction of Mechanics and Mathematics. Springer-Verlag, Berlin, 2009. Fundamentals.

[BL76] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[CO08] S. Conti and M. Ortiz. Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids*, 56:1885–1904, 2008.

[Dal93] G. Dal Maso. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[DF79] E. De Giorgi and T. Franzoni. On a type of variational convergence. In *Proceedings of the Brescia Mathematical Seminar, Vol. 3 (Italian)*, pages 63–101, Milan, 1979. Univ. Cattolica Sacro Cuore.

[Gho08] N. Ghoussoub. *Selfdual partial differential systems and their variational principles*. Universitext. Springer-Verlag, In press, 2008.

[Gio96] E. De Giorgi. Conjectures concerning some evolution problems. *Duke Math. J.*, 81(2):255–268, 1996.

[HLW06] E. Hairer, Ch. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[Ilm94] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108(520):x+90, 1994.

[Lán70] C. Lánczos. *The variational principles of mechanics*. Mathematical Expositions, No. 4. University of Toronto Press, Toronto, Ont., fourth edition, 1970.

[LM72] J.-L. Lions and E. Magenes. *Non-homogeneus boundary value problems and applications*, volume 1. Springer-Verlag, New York - Heidelberg, 1972.

[LS11] M. Liero and U Stefanelli. Elliptic regularization of hyperbolic-parabolic semilinear equations. *In preparation*, 2011.

[MO08] A. Mielke and M. Ortiz. A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems. *ESAIM Control Optim. Calc. Var.*, 14(3):494–516, 2008.

[Moi04] B. L. Moiseiwitsch. *Variational principles*. Dover Publications Inc., Mineola, NY, 2004. Corrected reprint of the 1966 original.

[MS08] A. Mielke and U. Stefanelli. A discrete variational principle for rate-independent evolution. *Adv. Calc. Var.*, 1(4):399–431, 2008.

[MS11] A. Mielke and U. Stefanelli. Weighted energy-dissipation functionals for gradient flows. *ESAIM Control Optim. Calc. Var.*, 17(1):52–85, 2011.

[RSSS11a] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli. Eliptic regularization of gradient flows in metric spaces. *In preparation*, 2011.

[RSSS11b] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli. A variational principle for gradient flows in metric spaces. *C. R. Math. Acad. Sci. Paris*, to appear, 2011.

[SS11] E. N. Spadaro and U. Stefanelli. A variational view at mean curvature evolution for linear growth functionals. *J. Evol. Equ.*, to appear, 2011.

[ST10] E. Serra and P. Tilli. Nonlinear wave equations as limits of convex minimization problems: proof of a conjecture by De Giorgi. *Preprint*, 2010.

[Ste11] U. Stefanelli. The De Giorgi conjecture on elliptic regularization. *Math. Models Meth. Appl. Sci.*, 21(6):1377–1394, 2011.