



DOUBLE MASTER'S DEGREE IN MATHEMATICS – MAPPA

---

UNIVERSITÉ PARIS DAUPHINE PSL  
CENTRE DE RECHERCHE EN MATHÉMATIQUES DE LA DÉCISION CEREMADE

UNIVERSITÀ DEGLI STUDI DI PADOVA  
DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

---

# Regularity for a class of rate-independent processes via weighted energy-dissipation principle

Giovanni Carlon

N° ÉTUDIANT: 22401609

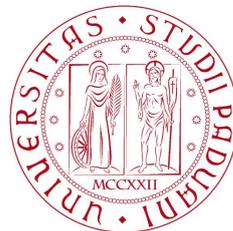
MATRICOLA: 2120841

SUPERVISOR:

Prof. Filippo Santambrogio

CO-SUPERVISOR:

Prof. Marco Cirant



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

---

ACADEMIC YEAR 2024/2025  
AUGUST 27, 2025

**Abstract:** This work studies rate-independent evolution problems in Banach spaces, which naturally generalize gradient flows via 1-homogeneous dissipation functionals. Such problems are governed by a doubly-nonlinear differential inclusion, whose solutions often exhibit low regularity due to the lack of super-linearity in the dissipation. However, under suitable uniform convexity and regularity assumptions on the energy functional, solutions can be shown to be Lipschitz continuous in time. The main contribution of this thesis is a new variational proof of this regularity result using the weighted energy-dissipation (WED) principle. This approach offers a purely variational alternative to existing proofs and extends recent results in more specific settings.

# Contents

<b>Introduction</b>	<b>5</b>
<b>1 Rate-independent processes</b>	<b>9</b>
1.1 Existence and uniqueness results . . . . .	12
1.1.1 Smooth and uniformly convex energies . . . . .	13
1.1.2 Nonconvex and nonsmooth problems . . . . .	14
1.2 Regularity . . . . .	15
1.2.1 Lipschitz regularity via uniform convexity . . . . .	16
1.2.2 Higher regularity . . . . .	16
1.3 Examples . . . . .	17
1.3.1 One dimensional examples . . . . .	17
1.3.2 Example with quadratic energy . . . . .	19
1.4 Weighted energy-dissipation principle . . . . .	20
<b>2 Original contributions</b>	<b>25</b>
<b>3 Proofs</b>	<b>29</b>
3.1 Proof of Theorem 2.1 . . . . .	29
3.1.1 Regularity in finite dimension . . . . .	29
3.1.2 Approximation of norms . . . . .	36
3.1.3 Regularity in infinite dimension . . . . .	41
3.2 Proof of Theorem (2.2) . . . . .	47
<b>References</b>	<b>51</b>



# Introduction

Rate-independent processes are among the most natural evolution problems for curves taking values in Banach spaces. They arise as a 1-homogeneous counterpart to the classical theory of gradient flows.

The simplest setting for this type of evolution problems consists in

- a Banach space  $X$ , which serves as the state space of the problem;
- a convex dissipation functional  $\Psi : X \rightarrow [0, +\infty]$ ;
- a time-dependent energy  $\mathcal{E} : [0, T] \times X \rightarrow [0, +\infty]$ .

Then, one seeks curves  $u : [0, T] \rightarrow X$ , with  $u(0) = u_0$  that follow the direction of maximal descending slope of the energy  $\mathcal{E}$ . This evolution can be characterized in terms of the dissipation  $\Psi$  and the differential  $D\mathcal{E}$  of the energy by the following doubly-nonlinear differential inclusion

$$0 \in \partial\Psi(\dot{u}(t)) + D\mathcal{E}(t, u(t)) \subset X^*, \quad \text{for } t \in (0, T) \quad (\text{SF})$$

In this framework, gradient flows correspond to super-linear dissipation functionals  $\Psi$ . The simplest example occurs when  $X$  is a Hilbert space and  $\Psi$  is its squared norm. In this case, (SF) takes the classical form

$$\dot{u}(t) = -\nabla\mathcal{E}(t, u(t)),$$

where one usually identifies  $X$  and  $X^*$  via the Riesz isomorphism.

On the other hand, rate-independent evolutions correspond to one homogeneous dissipations, so that (SF) is invariant under monotone reparametrizations of time. A typical example is when  $\Psi$  is the norm on the Banach space  $X$ . Although the formal structure of the equation appears similar to the gradient flow case, the loss of super-linearity introduces many new challenges. In particular, one usually obtains only BV estimates in time for the solutions, so that jumps can typically occur during the evolution, at least when the energy is not convex. Therefore, even for very regular energies and data, one has to deal with a non-smooth setting and possibly discontinuous solutions, which does not occur in the gradient flow case. However, when the energy  $\mathcal{E}$  is uniformly convex one can prove the following regularity result (see Section 1.2):

**Theorem.** Assume  $\mathcal{E}$  and  $\Psi$  to be defined on a Banach space  $(X, \|\cdot\|)$  and satisfy the following conditions:

1.  $\exists \alpha > 0, \forall t \in [0, T] : \mathcal{E}(t, \cdot)$  is  $\alpha$ -uniformly convex;
2.  $\exists C > 0, \forall t \in [0, T], \forall u_1, u_2 \in X : |\partial_t \mathcal{E}(t, u_1) - \partial_t \mathcal{E}(t, u_2)| \leq C \|u_1 - u_2\|$ .

Then, any solution  $u : [0, T] \rightarrow X$  is Lipschitz continuous and satisfies the estimate

$$\forall s, t \in [0, T] : \|u(t) - u(s)\| \leq \frac{C}{\alpha} |t - s|. \quad (1)$$

The main contribution of this work is a new proof of this regularity result by purely variational techniques. This is made possible by results obtained in [11], where the authors reformulate Problem (SF) in a weaker variational form. They introduce a *weighted energy-dissipation functional* (WED) of the form

$$\mathcal{I}_\lambda(u) = e^{-\lambda T} \mathcal{E}(T, u(T)) + \int_{[0, T]} e^{-\lambda t} \Psi(du) + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) dt$$

defined for  $u \in \text{BV}([0, T], X)$  with  $u(0) = u_0$  (here the first integral has to be interpreted in a Lebesgue-Steitjes sense), and prove that for  $\lambda \rightarrow +\infty$ , the minimizers of  $\mathcal{I}_\lambda$  converge to a solution of the rate-independent problem (SF); see Section 1.4 for the precise statement.

We prove that the minimizers of functionals of the form  $\mathcal{I}_\lambda$ , when  $\Psi$  is given by a norm and  $\mathcal{E}$  satisfies the hypothesis of the regularity theorem above, satisfy (1), and are therefore uniformly Lipschitz continuous in  $\lambda$ . This result draws on ideas originally introduced in [6], where the authors study a special case of our theorem. The uniform Lipschitz estimate, combined with the results of [11], allows us to establish the desired regularity of solutions to (SF).

This work is organized as follows. Chapter 1 provides a more detailed introduction to rate-independent processes, with a primary focus on the specific case considered in this study, namely, when the dissipation potential  $\Psi$  is given by a norm and the energy functional  $\mathcal{E}$  is smooth and uniformly convex. However, we also briefly outline the more general framework for these evolution problems, as developed in [8], [13], and related works. In this broader setting, one can allow for nonconvex and nonsmooth energy functionals, and the dissipation is modeled using a more general dissipation distance  $\mathcal{D}$ . In particular, we will present the energetic formulation of the problem, developed by A.Mielke and his collaborators [9, 10, 8], which allows one to construct solution via a time discretization scheme resembling De Giorgi's *minimizing movement* approximation scheme for gradient flows [1, 5]. Another key part of this chapter is Section (wed), where we describe the connection between the *weighted energy-dissipation functionals*  $\mathcal{I}_\lambda$

and the original rate-independent problem (SF).

Chapter 2 is devoted to the precise formulation and presentation of the main results of this work.

Finally, Chapter 3 is devoted entirely to the rigorous proofs of the results stated in the previous Chapter. Here, we develop the necessary analytical tools and carry out the variational arguments in detail.



# Chapter 1

## Rate-independent processes

Rate-independent evolutions arise naturally in several branches of applied mathematics, such as dry friction on surfaces, viscoelasticity and plasticity. Such systems respond to changes in external actions in a manner which is invariant for time rescalings, being driven by external loadings on a time scale much slower than their internal scale.

This kind of behavior is encoded in the simplest example of rate-independent evolution, namely the doubly nonlinear differential inclusion [4]

$$\begin{aligned}\partial\Psi(\dot{u}(t)) + D\mathcal{E}(t, u(t)) \ni 0, \\ u(0) = u_0 \in X,\end{aligned}\tag{SF}$$

where  $X$  is a Banach space;  $\Psi : X \rightarrow [0, +\infty]$  is a 1-homogeneous (i.e.,  $\Psi(\lambda v) = \lambda\Psi(v)$ , for every  $\lambda \geq 0$ ), convex and lower semi-continuous dissipation potential;  $\mathcal{E} : [0, T] \times X \rightarrow [0, +\infty]$  is an energy-storage functional and time  $t$  varies in  $[0, T]$  (here  $\partial$  denotes the subdifferential of a convex function and  $D$  denotes the gradient with respect to the second variable). Thus, the inclusion (SF) must be interpreted as a balance between dissipative and conservative forces.

The 1-homogeneity assumption on  $\Psi$  is the real source of rate-independence; indeed, it implies that for any strictly increasing diffeomorphism  $\alpha : [0, T] \rightarrow [0, \tilde{T}]$ , the rescaled function  $\tilde{u}(t) = u(\alpha(t))$  is a solution of (SF) with  $\tilde{\mathcal{E}}(t, \cdot) = \mathcal{E}(\alpha(t), \cdot)$ . This is the key distinction from classical rate-dependent gradient flows, where the dissipation depends on  $\dot{u}$  in a superlinear manner, thereby losing the property of time-rescaling invariance.

Classically, another way of modeling rate independent processes is through an evolutive variational inequality

$$\langle D\mathcal{E}(t, u(t)), v - \dot{u}(t) \rangle + \Psi(v) - \Psi(\dot{u}(t)) \geq 0, \forall t \in [0, T], \forall v \in X, \tag{VI}$$

here  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $X$  and  $X^*$ ; this formulation is completely equivalent to (SF) as one can immediately see from the definition of subd-

ifferential

$$\partial\Psi(w) = \{p \in X^* : \Psi(v) \geq \Psi(w) + \langle p, v - w \rangle, \forall v \in X\}$$

However, in recent years a new energetic approach has been developed to study the problem [8, 10, 9, 15, 12].

**Definition 1.1.** A curve  $u : [0, T] \rightarrow X$  is said to be an **energetic solution** of the rate independent system associated with  $(X, \mathcal{E}, \Psi)$ , if  $t \rightarrow \partial_t \mathcal{E}(t, u(t))$  is integrable and for all  $t \in [0, T]$  it satisfies the two following conditions:

$$\mathcal{E}(t, u(t)) \leq \mathcal{E}(t, \hat{u}) + \Psi(\hat{u} - u(t)), \quad \forall \hat{u} \in X; \quad (\text{S})$$

$$\mathcal{E}(t, u(t)) + \text{Diss}_\Psi(u, [0, t]) = \mathcal{E}(0, u(0)) + \int_0^t \partial_s \mathcal{E}(s, u(s)) \, ds \quad (\text{E})$$

In the above definition  $\text{Diss}_\Psi(u, [0, t])$  denotes the total dissipation of the curve  $u$  on the interval  $[0, t]$ , which for an arbitrary interval  $J$  is defined as

$$\text{Diss}_\Psi(u, J) = \sup \left\{ \sum_{k=1}^N \Psi(u(t_k) - u(t_{k-1})) : N \in \mathbb{N}, t_0, t_N \in J, t_0 < \dots < t_N \right\}$$

The first condition (S) is a global stability condition, and can be reformulated in terms of sets of stable states

$$\mathcal{S}(t) = \{u \in X : \mathcal{E}(t, u) \leq \mathcal{E}(t, \hat{u}) + \Psi(\hat{u} - u), \forall \hat{u} \in X\},$$

by just saying  $u(t) \in \mathcal{S}(t)$  for all  $t \in [0, T]$ . Observe that (S) means that  $u(t)$  is a global minimizer of  $\hat{u} \rightarrow \mathcal{E}(t, \hat{u}) + \Psi(\hat{u} - u(t))$ ; its immediate interpretation is that the solution  $u$  minimizes the energy  $\mathcal{E}$  while keeping into account the dissipated energy. The second condition (E) is a global energy balance between final, dissipated, initial energy and the work done by external forces on any interval  $[0, t]$ .

Observe that this definition of solution implies the two natural requirements for evolutionary problems, namely that restrictions and concatenation of solutions remain solutions. To be more precise, for any solution  $u : [0, T] \rightarrow X$  and any subinterval  $[s, t] \subseteq [0, T]$ , the restriction  $u|_{[s, t]}$  is an energetic solution with initial value  $u(s)$ . Moreover, if  $u_1 : [0, t] \rightarrow X$  and  $u_2 : [t, T] \rightarrow X$  are solutions on their respective time intervals and if  $u_1(t) = u_2(t)$ , then the concatenation of the two curves  $u : [0, T] \rightarrow X$  is an energetic solution over  $[0, T]$ .

For general energies, this new notion of solution does not coincide with (SF) and (VI); however, if the energy  $\mathcal{E}$  is convex, one can prove that they are in fact equivalent (see [10]). The intuition behind this equivalence for convex energies

could be the following: for a general smooth curve  $u : [0, T] \rightarrow X$ , by Fenchel-Young inequality one has

$$\frac{d}{dt}\mathcal{E}(t, u(t)) - \partial_t \mathcal{E}(t, u(t)) = \langle D\mathcal{E}(t, u(t)), \dot{y}(t) \rangle \geq -\Psi(\dot{u}(t)) - \Psi^*(-D\mathcal{E}(t, u(t))),$$

where  $\Psi$  denotes the Legendre-Fenchel transform of the convex dissipation potential  $\Psi$ . Integrating the previous inequality over the time interval  $[0, t]$  we obtain

$$\mathcal{E}(t, u(t)) + \int_0^t [\Psi(\dot{u}(s)) + \Psi^*(-D\mathcal{E}(s, u(s)))] ds \geq \mathcal{E}(0, u(0)) + \int_0^t \partial_s \mathcal{E}(s, u(s)) ds.$$

Now, under our hypotheses of convexity and 1-homogeneity of  $\Psi$ , one can prove that  $\Psi$  is the support function of a closed convex set  $F^* \subset X^*$  and thus  $\Psi^* = \chi_{F^*}$ , the indicator function (in the sense of convex analysis) of  $F^*$ , and  $\text{Diss}_\Psi(u, [0, t]) = \int_0^t \Psi(\dot{u}(s)) ds$ , the total dissipated energy. Thus the previous identity is equivalent to the following two conditions

$$\begin{cases} -D\mathcal{E}(s, u(s)) \in F^*, & \forall s \in [0, t], \\ \mathcal{E}(t, u(t)) + \text{Diss}_\Psi(u, [0, t]) \geq \mathcal{E}(0, u(0)) + \int_0^t \partial_s \mathcal{E}(s, u(s)) ds. \end{cases}$$

One can easily see that for convex energies the stable sets are implicitly defined by  $\mathcal{S}(t) = \{u : -D\mathcal{E}(s, u) \in F^*\}$ ; so that, the first condition means  $u(s) \in \mathcal{S}(s)$  for all  $s \in [0, t]$ , which is (S) in Definition 1.1.

Meanwhile, the second condition is (E) provided one has equality instead of inequality, this happens precisely when one has equality at all times in the Fenchel-Young inequality we used above, that is, if and only if  $\partial\Psi(\dot{u}(t)) \in -D\mathcal{E}(t, u(t))$ , for all  $t \in [0, T]$ , which is the original subdifferential inclusion (SF).

Obviously, this is only an heuristic argument for general  $\Psi$  and  $\mathcal{E}$ , since, normally, a solution of (S)&(E) does not possess a time derivative, and thus one has to be more careful with any term involving  $\dot{u}$ .

The major advantage of this energetic formulation is that it does not require any derivative of  $\Psi$ ,  $\mathcal{E}$  or  $u$ , except for the time derivative of  $t \rightarrow \mathcal{E}(t, \hat{u})$ . Thus, one is able to define a notion of solution even when  $\mathcal{E}(t, \cdot)$  is non-differentiable, allowing the theory describe problems in mechanics where the energies of the systems have very low regularity. However, in this case one has to allow the solutions to have discontinuous trajectories.

This chapter will be organized as follows. In Section 1.1 we will introduce some well-posedness results for the energetic formulation given above. We will treat two cases: the first consists of systems with smooth and uniformly convex energies defined on a Banach space, the second is the one of nonconvex and nonsmooth energies; in this last case the theory can be generalized to treat situations where the

state space does not possess a linear structure and the dissipation is implemented through a more general dissipation distance  $\mathcal{D}$ , however, we will not treat this case in detail and we will just give a result for nonconvex energies over a Banach space, as this is setting of our regularity results of Chapter 2.

In Section 1.2 we will show how one can obtain regularity for the solutions of the energetic formulation under some additional hypotheses on the energy.

Finally, in Section 1.3, we will present some examples of rate-independent systems for which the results of the previous sections can be applied. In particular, we will show that even when the state space is one dimensional, the solutions of the energetic formulation may have discontinuous trajectories.

Most of the results of this Chapter will be presented without proof; this is a choice we took in order to make the presentation of the topic lighter, and mostly because the techniques we used in the proof of our results of Chapter 2 are completely different from the ones used to prove the result we are going to present in this Chapter. Nevertheless, we refer the interested reader to the very complete survey [8], which is where most of the results of this Chapter are taken from, and to [13] for a general overview of the possible applications of rate-independent processes in mechanics.

## 1.1 Existence and uniqueness results

Most existence results in the theory of rate-independent processes were obtained for the energetic formulation of Definition 1.1; this is due to the fact that this is the only notion of solution that generalizes to more general nonlinear and nonconvex situations like the one we will briefly introduce in Section 1.1.2, and that there exists a natural approach to solve (S)&(E) via time discretization.

For this, fix a sequence of partitions  $0 = t_0^N < t_1^N < \dots < t_{N-1}^N < t_N^N = T$  of the interval  $[0, T]$  depending on the parameter  $N \in \mathbb{N}$ , such that the finiteness  $\Delta_N = \max\{t_k^N - t_{k-1}^N : k = 1, \dots, N\}$  tends to zero as  $N \rightarrow +\infty$ . We solve iteratively for a given initial value  $u_0 \in \mathcal{S}(0)$  and each  $N$  the minimization problem

$$u_k^N \in \arg \min\{\mathcal{E}(t_k^N, u) + \Psi(u - u_{k-1}^N) : u \in X\}$$

Then, define the piecewise constant function  $u^N$  taking the values  $u_k^N$  on the interval  $[t_{k-1}^N, t_k^N]$  for each  $k = 1, \dots, N$ . To conclude now, one just needs to show that the family  $(u^N)_N$  is precompact in some appropriate topology and the limit function satisfies the energetic formulation (S)&(E).

The existence results of the following sections were obtained in [10] and [8] using this approach.

On the other hand, most uniqueness results are obtained by working directly on

the definition of energetic solution. In general, uniqueness of solutions holds only under the hypotheses of strict convexity of the energy  $\mathcal{E}$ .

The second approach is based on the theory of set-valued monotone operators and heavily relies on an underlying Hilbert or Banach structure. One can also treat nonsmooth problems by means of regularization techniques. We refer to [4] for this approach, where some existence results for doubly nonlinear differential inclusions were obtained using the Yosida regularization of a monotone operator.

### 1.1.1 Smooth and uniformly convex energies

In this section we treat the case of general smooth and uniformly convex energies  $\mathcal{E}$  on a reflexive Banach space  $(X, \|\cdot\|)$ . This case was thoroughly analyzed in [8], which was inspired by the previous works [9, 15, 10].

The dissipation potential  $\Psi : X \rightarrow [0, +\infty]$  is assumed to be convex, 1-homogeneous and lower semicontinuous; recall that this implies that  $\Psi$  is the Legendre-Fenchel transform of the indicator function of the closed convex set  $F^* = \partial\Psi(0) \subset X^*$ . The value  $+\infty$  is allowed, as well as  $\Psi(v) = 0$  even if  $v \neq 0$ . Importantly, contrary to [10], we do not require the assumption  $\rho_1\|v\| \leq \Psi(v) \leq \rho_2\|v\|$ , for  $0 < \rho_1 \leq \rho_2$ . Thus, this section can be used to model various situation in continuum mechanics where  $X = L^2(\Omega)$  and  $\Psi(v) = \|v\|_{L^1(\Omega)}$ .

Throughout this section we will assume  $\mathcal{E} \in C^3([0, T] \times X, [0, +\infty])$  and the existence of a positive constant  $C_E > 0$  such that

$$\mathcal{E}(t, u) < \infty \implies \begin{cases} \partial_t \mathcal{E}(\cdot, u) : [0, T] \rightarrow \mathbb{R} \text{ is measurable} \\ \text{and } |\partial_t \mathcal{E}(t, u)| \leq C_E(\mathcal{E}(t, u) + 1); \end{cases}$$

and that for each  $E_0 > 0$  there exist constants  $C_{ty}, C_{yyy}, C_{tyy} > 0$  such that

$$\mathcal{E}(t, y) < E_0 \implies \begin{cases} \|\mathbb{D}\mathcal{E}(t, y)\|, \|\mathbb{D}^2\mathcal{E}(t, y)\|, \|\partial_t \mathbb{D}\mathcal{E}(t, y)\| \leq C_{ty}, \\ \|\mathbb{D}^3\mathcal{E}(t, y)\| \leq C_{yyy}, \|\partial_t \mathbb{D}^2\mathcal{E}(t, y)\| \leq C_{tyy}, \end{cases}$$

where the norms have to be intended as the appropriate operator norms. But most importantly, we will make the hypotheses of uniform convexity of the energy, that is  $\exists \alpha > 0$  such that

$$\mathcal{E}(t, (1 - \theta)u_1 + \theta u_2) \leq (1 - \theta)\mathcal{E}(t, u_1) + \theta\mathcal{E}(t, u_2) - \frac{\alpha}{2}\theta(1 - \theta)\|u_1 - u_2\|^2, \quad (1.1)$$

for all  $\theta \in [0, 1]$  and  $u_1, u_2 \in X$ ; we recall that because of  $C^2$ -regularity, the condition above is equivalent to

$$\exists \alpha > 0, \forall u, v \in X : \langle \mathbb{D}^2\mathcal{E}(t, u)v, v \rangle \geq \alpha\|v\|^2;$$

Because of strict convexity of the energy over a Banach space  $X$ , the energetic formulation (S)&(E) is equivalent to (SF) and (VI) (see for instance [10], Theorems 3.4 & 3.5, for a rigorous proof of this fact).

In the uniformly convex and smooth case one can also fairly easily obtain stability of the solution with respect to initial data.

**Proposition 1.1.1** ([8], Proposition 4.1). *Let  $u_1, u_2 : [0, T] \rightarrow X$  be two solutions of (VI). Then there exists a constant  $K > 0$  such that for all  $t \in [0, T]$  one has*

$$\|u_1(t) - u_2(t)\| \leq K \|u_1(0) - u_2(0)\|$$

But most importantly, one can prove well-posedness of the problem.

**Theorem 1.2** ([8], Theorem 4.3). *For each  $u_0 \in \mathcal{S}(0)$  there exists a unique solution of  $u \in W^{1,\infty}([0, T], X)$  of (S)&(E) with  $u(0) = u_0$ .*

The existence part of this Theorem was obtained by the authors by means of the time discretization scheme described above. In particular, the reflexivity hypotheses on  $X$  is only used to obtain the necessary precompactness of the piecewise constant solutions of the incremental problem, and can be replaced by some weaker compactness hypotheses of the sublevels of  $\mathcal{E}$ . In this very simple case the convergence of the discretized solutions to the limit is in the  $L^\infty([0, T], X)$  norm. The regularity of the solution is a consequence of the uniform convexity of  $\mathcal{E}(t, \cdot)$  (see in Theorem 1.4 below), as well as the Lipschitz stability with respect to initial data (which would give another proof of the uniqueness of solutions).

## 1.1.2 Nonconvex and nonsmooth problems

The existence theory for nonconvex and nonsmooth energies is often formulated in a more general nonlinear setting [8, 7], in order to be able to treat a wider range of mechanical application. Usually, the state space is chosen as the set  $\mathcal{X} = \mathcal{F} \times \mathcal{Z}$ ; this splitting is chosen such that the dissipation of the system only affects the variables  $z \in \mathcal{Z}$ , and has no effect on the so called elastic variables  $\varphi \in \mathcal{F}$ . Moreover, the dissipation is implemented through a more general dissipation distance  $\mathcal{D} : \mathcal{Z} \times \mathcal{Z} \rightarrow [0, +\infty]$ , which generalizes  $\Psi$  via  $\mathcal{D}(z_1, z_2) = \Psi(z_2 - z_1)$ , and that does not require a linear structure on  $\mathcal{Z}$ .

In this very general nonlinear setting, the formulations (SF)&(VI) do not make sense; however, the notion of solution via (S)&(E) is very easily generalized by just substituting  $\Psi(\hat{u} - u(t))$  by  $\mathcal{D}(u(t), \hat{u})$  in (S), and by replacing  $\text{Diss}_\Psi(u, [0, t])$  by the new total dissipation

$$\text{Diss}_\mathcal{D}(u, [0, t]) = \sup \left\{ \sum_{k=1}^N \mathcal{D}(u(t_{k-1}), u(t_k)) : N \in \mathbb{N}, 0 = t_0 < \dots < t_N = t \right\}.$$

Furthermore, one can obtain existence for the energetic formulation by a nonlinear analogue of the incremental minimization scheme of the previous section.

The complete treatment of this case goes beyond the scope of this thesis, since our regularity results of Chapter 3 are derived with an underlying linear structure and with a uniform convexity assumptions on the energy of the problem, and thus will mostly apply to the systems described in the previous section. However, we include an existence result for nonconvex and nonsmooth energies defined on a nonreflexive Banach space, which can be obtained as an easy consequence of the nonlinear theory just described.

**Theorem 1.3** ([8], Theorem 5.8). *Let  $Z$  and  $X$  be Banach spaces. Suppose that  $X$  is compactly embedded in  $Z$  and that  $B_X$  is closed in  $Z$ , and that  $\Psi(\cdot) = \|\cdot\|_Z$ . Furthermore, the functional  $\mathcal{E} : [0, T] \times Z \rightarrow [0, +\infty]$  has the following properties:*

1.  $\mathcal{E}$  is lower semi-continuous on  $[0, T] \times Z$  (with respect to the strong topology of  $Z$ );
2. For some real numbers  $c_1 > 0$ ,  $C_2$  and  $\alpha > 0$  we have

$$\mathcal{E}(t, x) \geq c_1 \|x\|_H^\alpha - C_2;$$

3. There exists  $C_E > 0$  such that for all  $x \in X$ :

$$\begin{aligned} \mathcal{E}(t, x) < \infty &\implies \partial_t \mathcal{E}(\cdot, x) : [0, T] \rightarrow R \text{ is measurable} \\ &\text{and } |\partial_t \mathcal{E}(t, x)| \leq C_E(\mathcal{E}(t, x) + 1); \end{aligned}$$

4. For all  $E_* > 0$  and  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$\mathcal{E}(t, x_*) \leq E_* \text{ and } |t - s| \leq \delta \implies |\partial_t \mathcal{E}(t, x_*) - \partial_t \mathcal{E}(s, x^*)| < \varepsilon.$$

Then, for each  $x_0 \in \mathcal{S}(0)$  there exists at least one solution  $x \in \text{BV}([0, T], Z) \cap L^\infty([0, T], X)$  of (S)&(E) with  $x(0) = x_0$ .

We will directly apply this result in the next Chapter, when we will talk about the Weighted energy-dissipation principle and its application to the existence theory for rate-independent processes.

## 1.2 Regularity

As we already described at the start of this Chapter, in general, a solution of the differential inclusion (SF) will have jumps in its trajectory, and thus, will not be continuous. However, when the energy functional  $\mathcal{E}$  possesses some convexity properties, one can show that the solutions of (SF) are continuous. This will be the topic of this section.

### 1.2.1 Lipschitz regularity via uniform convexity

**Theorem 1.4** ([8], Theorem 3.4). *Assume  $\mathcal{E}$  and  $\Psi$  to be defined on a Banach space  $(X, \|\cdot\|)$  and satisfy the following conditions:*

1.  $\exists \alpha > 0, \forall t \in [0, T] : \mathcal{E}(t, \cdot)$  is  $\alpha$ -uniformly convex (in the sense of (1.1));
2.  $\exists C > 0, \forall t \in [0, T], \forall u_1, u_2 \in X : |\partial_t \mathcal{E}(t, u_1) - \partial_t \mathcal{E}(t, u_2)| \leq C \|u_1 - u_2\|$ .

*Then, any energetic solution  $u : [0, T] \rightarrow X$  is Lipschitz continuous and satisfies the estimate*

$$\forall s, t \in [0, T] : \|u(t) - u(s)\| \leq \frac{C}{\alpha} |t - s|. \quad (1.2)$$

*Proof.* Let us first observe that, thanks to the first hypotheses and the convexity of  $\Psi$ , we have that  $\forall (t, u) \in [0, T] \times X$ , the function  $\hat{u} \rightarrow \mathcal{E}(t, \hat{u}) + \Psi(\hat{u} - u)$  is  $\alpha$ -uniformly convex with respect to  $\|\cdot\|$ . We recall that for the minimizer  $y_*$  of an  $\alpha$ -uniformly convex function  $f$  one always has  $f(y) \geq f(y_*) + \frac{\alpha}{2} \|y - y_*\|^2$ , which in particular implies that the minimizer is unique. But now, using (S), we have that  $u(s)$  is the minimizer of  $\hat{u} \rightarrow \mathcal{E}(s, \hat{u}) + \Psi(\hat{u} - u(s))$ , so that

$$\begin{aligned} \frac{\alpha}{2} \|u(s) - u(t)\|^2 &\leq \mathcal{E}(s, u(t)) + \Psi(u(t) - u(s)) - \mathcal{E}(s, u(s)) \\ &\leq \mathcal{E}(t, u(t)) + \text{Diss}_\Psi(u, [s, t]) - \mathcal{E}(s, u(s)) - \int_s^t \partial_\tau \mathcal{E}(\tau, u(t)) \, d\tau \\ &\leq \int_s^t [\partial_\tau \mathcal{E}(\tau, u(\tau)) - \partial_\tau \mathcal{E}(\tau, u(t))] \, d\tau \leq C \int_s^t \|u(\tau) - u(t)\| \, d\tau \end{aligned}$$

Where we used  $\Psi(u(t) - u(s)) \leq \text{Diss}_\Psi(u, [s, t])$  and (E). Now, for fixed  $t$  let  $\delta(s) = \|u(s) - u(t)\|$ ; then we have just shown that  $\delta(s)^2 \leq \frac{2C}{\alpha} \int_s^t \delta(\tau) \, d\tau$  for  $s \in [0, t]$ . Let  $h(r) = \int_{t-r}^t \delta(\tau) \, d\tau$  for  $r \geq 0$ ; then  $h(0) = 0$  and  $h'(r) \leq (2Ch(r)/\alpha)^{1/2}$ , which implies  $h(r) \leq Cr^2/(2\alpha)$  and then  $\delta(t-r) \leq Cr/\alpha$ , which is the desired result.  $\square$

This result directly applies to the cases of Section 1.1.1, however we must point out that one can obtain an analogous result for rate-independent processes defined over a metric space  $(\mathcal{X}, d)$  and where the uniform convexity and the Lipschitz continuity of  $\partial_t \mathcal{E}(t, \cdot)$  are with respect to the metric  $d$ . Thus, this result can also be applied to the cases described at the start of Section 1.1.2.

### 1.2.2 Higher regularity

Under some additional hypotheses on the regularity of the solutions can be improved from the Lipschitz continuity of Theorem 1.4 to  $\dot{y} \in \text{BV}([0, T], X)$ , but this is the best we can hope for (see Example 1.7).

**Theorem 1.5.** *Assume in addition to the hypotheses of Theorem 1.4, that  $\mathcal{E} \in C^3([0, T] \times X, \mathbb{R})$  and  $\Psi = (I_{F^*})^*$ , where  $F^*$  satisfies  $B_{\rho_1}^*(0) \subseteq F^* \subseteq B_{\rho_2}^*(0)$ , or equivalently  $\rho_1 \|\cdot\| \leq \Psi(\cdot) \leq \rho_2 \|\cdot\|$ , for some constants  $0 < \rho_1 < \rho_2$ , and that its boundary  $\partial F^*$  is of class  $C^2$ , then any solution of (VI) satisfies  $\dot{y} \in BV([0, T], X)$ .*

This result is proved in [10], Theorem 7.8. This higher temporal regularity can be used to improve the rate of convergence of the solutions to the incremental problem; in addition one can also obtain convergence of the time derivative in the  $L^2([0, T], X)$  norm, if one considers the piecewise linear interpolants instead of the piecewise constant ones (see [8], Theorem 4.11).

## 1.3 Examples

In this section, we present several examples of rate-independent systems to illustrate the theoretical concepts discussed earlier. These examples are chosen to highlight key features such as hysteresis and non-smooth behavior. Wherever possible, we also demonstrate the systems' time evolution to provide a more intuitive understanding of their dynamics.

### 1.3.1 One dimensional examples

The simplest examples of rate independent evolution are formulated over the real line  $X = \mathbb{R}$ , with the dissipation potential given by  $\Psi(v) = \kappa|v|$ . Although this setting may appear overly simplistic, it already captures the rather different behavior of the solutions when transitioning from convex to nonconvex energies. The major advantage of this one dimensional case is that solutions can be computed explicitly fairly easily directly from the definition.

*Example 1.6* ([10], Example 7.7). Consider the rate-independent system with  $X = \mathbb{R}$ ,  $\Psi(v) = |v|$  and  $\mathcal{E}_1(t, u) = \frac{1}{2}u^2 - \ell(t)u$  with  $\ell(t) = 4t - t^2$ . This system falls under the cases investigated in Section 1.1. So for every stable initial condition there exists a unique Lipschitz continuous solution, thanks to Theorem 1.2. In particular for the initial condition  $u(0) = 0 \in [-1, 1] = \mathcal{S}(0)$ , some easy computations give

$$u(t) = \begin{cases} 0 & \text{for } t \leq 2 - \sqrt{3}, \\ 4t - t^2 - 1 & \text{for } t \in [2 - \sqrt{3}, 2], \\ 3 & \text{for } t \in [2, 2 + \sqrt{2}], \\ 4t - t^2 + 1 & \text{for } t \geq 2 + \sqrt{2}. \end{cases}$$

A nice visual representation of the solution is given in the  $(u, \ell)$ -plane (see Figure 1.1). In particular, one can easily observe that  $u(t) \in \mathcal{S}(t) = [\ell(t) - 1, \ell(t) + 1]$  at

every time  $t$ . Furthermore, we observe that  $\dot{u}(t) \in \text{BV}([0, T])$ , in accordance with Theorem 1.5.

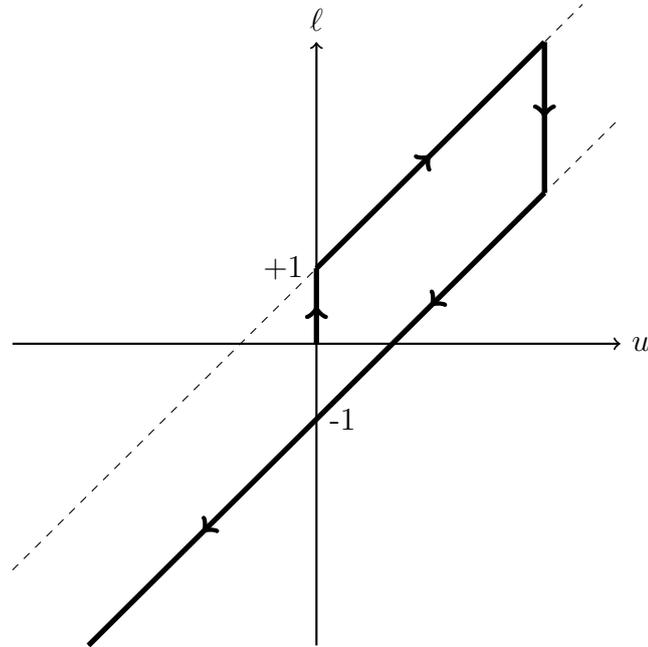


Figure 1.1: Plot of the solution (thick line) in the  $(u, \ell)$  plane. The arrows indicate the positive direction of time

*Example 1.7* ([11], Example 3.2). Consider the rate-independent system with  $X = \mathbb{R}$ ,  $\Psi(v) = \frac{1}{2}|v|$  and  $\mathcal{E}_2(t, u) = F(u) - \ell(t)u$ . Where  $F$  is a double-well potential, defined as

$$F(u) = \begin{cases} \frac{1}{2}(u+1)^2 & \text{for } u \leq -\frac{1}{2}, \\ \frac{1}{4} - \frac{1}{2}u^2 & \text{for } |u| \leq \frac{1}{2}, \\ \frac{1}{2}(u-1)^2 & \text{for } u \geq \frac{1}{2}. \end{cases}$$

In this case the energy is not convex; nevertheless, one can still obtain existence and uniqueness for (S)&(E), however the solution obtained will be discontinuous. In the special case  $\ell(t) = \min\{t - 1, 5 - t\}$  with initial condition  $u(0) = -2$ , the

solution on the interval  $[0, 7]$  can be recovered easily from the energy balance (E):

$$u(t) = \begin{cases} -2 & \text{for } t \in [0, 1/2], \\ t - 5/2 & \text{for } t \in [1/2, 3/2), \\ t - 1/2 & \text{for } t \in (3/2, 3], \\ 5/2 & \text{for } t \in [3, 4], \\ 13/2 - t & \text{for } t \in [4, 11/2), \\ 9/2 - t & \text{for } t \in (11/2, 7]. \end{cases}$$

One can observe that, as often happens with nonconvex energies, the solution has two discontinuities at times  $t = 3/2$  and  $t = 11/2$  (see Figure 1.2).

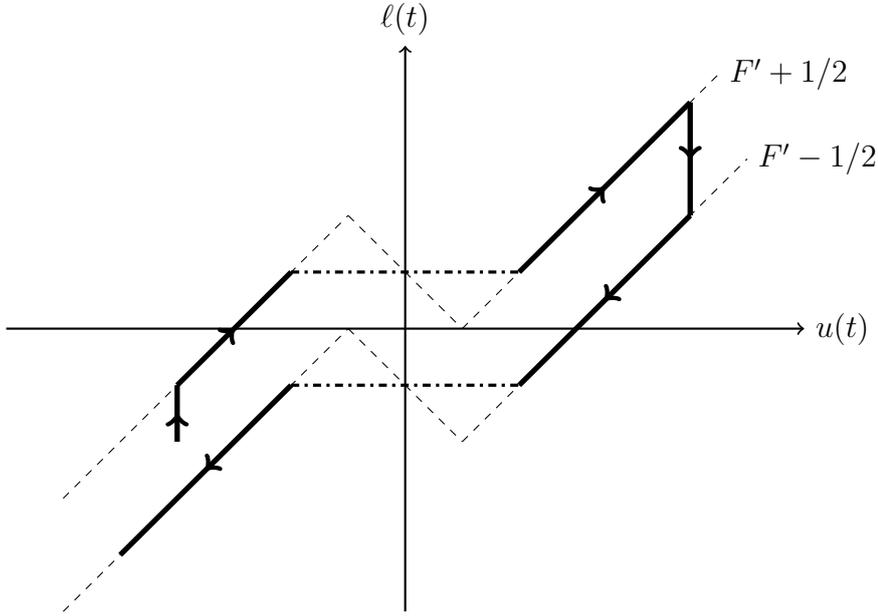


Figure 1.2: Plot of the solution in the  $(u, \ell)$ -plane

### 1.3.2 Example with quadratic energy

A typical situation in continuum mechanical problems is that the state variable  $x \in X$  consists of two components; an elastic component  $u \in U$  and a dissipative component  $z \in Z$ . The splitting is such that the dissipation functional  $\Psi : Y \rightarrow [0, +\infty]$  depends only on the  $\dot{z}$  component of  $\dot{y}$ . In particular we have

$$x = (u, z) \in U \times Z = X, \quad \Psi(\dot{x}) = \Psi((\dot{u}, \dot{z})) = \tilde{\Psi}(\dot{z})$$

with  $\tilde{\Psi}(\dot{z}) > 0$  when  $\dot{z} \neq 0$ . A possible choice for the energy functional  $\mathcal{E}$  is the quadratic one, i.e.  $\mathcal{E}$  and takes the form  $\mathcal{E}(t, x) = \frac{1}{2}\langle Ay, y \rangle - \langle \ell(t), y \rangle$ , where the linear operator  $A$  is given by

$$A = \begin{pmatrix} A_{UU} & A_{UZ} \\ A_{ZU} & A_{ZZ} \end{pmatrix} \in \text{Lin}(U \times Z, U^* \times Z^*), \quad A_{UU} > 0$$

and  $\ell = (\ell_U, \ell_Z) \in C([0, T], U^* \times Z^*)$  plays the role of an external loading in our system. With these choices, the subdifferential formulation takes the form

$$0 = A_{UU}u + A_{UZ}z - \ell_U(t) \in U^*, \quad \text{and} \quad \in \partial\tilde{\Psi}(\dot{z}) + A_{ZU}u + A_{ZZ}z - \ell_Z(t) \in Z^*.$$

The first equation is often called elastic equilibrium condition and can be solved uniquely for  $u$  given  $z$  and  $\ell_U$ , while the second inclusion is the flow rule of the rate-independent internal variable  $z$ .

## 1.4 Weighted energy-dissipation principle

In [11], the authors propose a new method for obtaining solutions to (SF) by what they call *weighted energy-dissipation principle*. With this method the solution of the rate-independent system with initial datum  $u(0) = u_0$  is obtained as the limit of minimizers to the parametrized family of functionals  $(\mathcal{I}_\lambda)_{\lambda>0}$ , where

$$\mathcal{I}_\lambda(u) = e^{-\lambda T} \mathcal{E}(T, u(T)) + \int_{[0, T]} e^{-\lambda t} \Psi(\mathrm{d}u) + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) \, dt$$

is defined for curves  $u \in \mathcal{Y}$  with  $u(0) = u_0$  ( $\mathcal{Y}$  denotes some suitable space of paths to be specified after making some assumptions on  $\Psi$  and  $\mathcal{E}$ ). The first integral represents a weighted total dissipation, which can be defined as

$$\sup \left\{ \sum_{k=0}^{N-1} e^{-\lambda t_{k+1}} \Psi(u(t_{k+1}) - u(t_k)) : 0 = t_0 < \dots < t_N = T \right\},$$

while the second integral corresponds to a weighted total energy.

Let us present some of the intuition behind this choice of functional. As we already explained in Section 1.1, a common way of obtaining solutions of (SF) is through an incremental minimization scheme based on time discretization. In particular, recall that the incremental problem presented in Section 1.1 consisted of a sequence of minimum problems

$$\inf_{u_{k+1} \in X} F_{k+1}(u_{k+1}, u_k), \quad k = 0, \dots, N-1 \tag{1.3}$$

where

$$F_{k+1}(u_{k+1}; u_k) = \Delta t \Psi \left( \frac{u_{k+1} - u_k}{\Delta t} \right) + \mathcal{E}(t_{k+1}, u_{k+1}) - \mathcal{E}(t_k, u_k), \quad (1.4)$$

for  $t_k = k\Delta t$ , with  $\Delta t > 0$ , and  $u_0 \in \mathcal{S}(0)$ .

Note that the small time discretization parameter  $\Delta t$  has no influence in the minimization (1.3) due to the 1-homogeneity of  $\Psi$ , but we still include the extra term  $-\mathcal{E}(t_k, u_k)$  in order for the dissipative and energetic terms to have the same order in  $\Delta t$ . In addition, we recall that the problems (1.3) are to be solved *incrementally/causally*, that is: problem  $k = 0$  is solved first, with initial datum  $u_0$ , to compute  $u_1$ ; then problem  $k = 1$  is solved, knowing the value of  $u_1$ , in order to compute  $u_2$ ; and so on.

This sequence of incremental problems can be collected into a single minimum problem on the entire discretized trajectory  $u = \{u_1, \dots, u_N\}$  using the concept of *Pareto optimality*. A candidate for such a functional to minimize could be

$$\mathcal{I}(u; \omega) = \sum_{k=0}^{N-1} \omega_{k+1} F_{k+1}(u_{k+1}; u_k), \quad (1.5)$$

where  $\omega = \{\omega_1, \dots, \omega_N\}$  are the so called positive *Pareto weights*. To ensure causality one can consider a sequence of weights  $\omega_1^\lambda \gg \dots \gg \omega_N^\lambda$ , indexed by a parameter  $\lambda$  to satisfy

$$\lim_{\lambda \rightarrow +\infty} \frac{\omega_{k+1}^\lambda}{\omega_k^\lambda} = 0$$

With this choice, as  $\lambda \rightarrow +\infty$ , the minimization of  $\mathcal{I}$  gives disproportionately larger importance to the first minimization problem relative to the second, to the second relative to the third, and so on.

If in addition we suppose  $\omega_k^\lambda = \omega_\lambda(t_k)$  for some function  $\omega_\lambda : [0, T] \rightarrow (0, +\infty)$  such that

$$\lim_{\lambda \rightarrow +\infty} \frac{\omega_\lambda(t)}{\omega_\lambda(s)} = 0, \quad \forall s, t \in [0, T], s < t$$

then causality is preserved and (1.5) becomes

$$\mathcal{I}(u; \omega_\lambda) = \sum_{k=0}^{N-1} \omega^\lambda(t_{k+1}) \left\{ \Psi \left( \frac{u_{k+1} - u_k}{\Delta t} \right) + \frac{\mathcal{E}(t_{k+1}, u_{k+1}) - \mathcal{E}(t_k, u_k)}{\Delta t} \right\} \Delta t,$$

which can be regarded as a time discretization of the following time-continuous functional

$$\begin{aligned} \tilde{\mathcal{I}}_\lambda(u) &= \int_0^T \omega_\lambda \left[ \Psi(\dot{u}) + \frac{d}{dt} \mathcal{E}(t, u) \right] dt \\ &= \omega_\lambda(T) \mathcal{E}(T, u(T)) - \omega_\lambda(0) \mathcal{E}(0, u(0)) + \int_0^T [\omega_\lambda \Psi(\dot{u}) - \dot{\omega}_\lambda \mathcal{E}(t, u)] dt, \end{aligned}$$

where in the last step we have performed an integration by parts.

An admissible and simple choice is  $\omega_\lambda(t) = e^{-\lambda t}$ , from which we recover the weighted energy dissipation functional presented above, up to the constant term  $\mathcal{E}(0, u(0)) = \mathcal{E}(0, u_0)$ .

Another connection between the weighted energy-dissipation functionals  $\mathcal{I}_\lambda$  and (SF) is that formally the Euler-Lagrange system for the functional  $\mathcal{I}_\lambda$  consists of

$$\begin{cases} \mathrm{D}\Psi(\dot{u}) + \mathrm{D}\mathcal{E}(t, u) - \frac{1}{\lambda}\mathrm{D}^2\Psi(\dot{u})\ddot{u} = 0, \\ u(0) = u_0, \\ \mathrm{D}\Psi(\dot{u}(T)) + \mathrm{D}\mathcal{E}(T, u(T)) = 0, \end{cases} \quad (1.6)$$

which can be interpreted as an *elliptic regularization* of the original differential inclusion (SF). One may therefore expect that minimizers converge to solutions of the equation  $\mathrm{D}\Psi(\dot{u}) + \mathrm{D}\mathcal{E}(t, u) = 0$ .

Let us now introduce the assumptions required in order to rigorously prove the convergence of minimizers of  $\mathcal{I}_\lambda$  to solutions of the rate-independent system. Let  $(X, \|\cdot\|_X)$  be a separable, reflexive Banach space that is compactly embedded into a larger Banach space  $(Z, \|\cdot\|_Z)$ . Assume  $\Psi : Z \rightarrow [0, +\infty]$  to be convex, 1-homogeneous, lower semi-continuous on  $X$  and such that

$$\exists c > 0, \forall v \in Z : \quad \Psi(v) \geq c\|v\|_Z, \quad (1.7)$$

and let  $\mathcal{E} : [0, T] \times Z \rightarrow \mathbb{R}_\infty$  be weakly lower semicontinuous and satisfying

$$\exists \alpha, c_1, C_2 > 0, \forall (t, v) \in Z : \quad \mathcal{E}(t, v) \geq c_1\|v\|_X^\alpha - C_2, \quad (1.8)$$

and there exists  $C_\mathcal{E} > 0$  such that for all  $v \in Z$ ,  $\partial_t \mathcal{E}(\cdot, v) : [0, T] \rightarrow \mathbb{R}$  is measurable and

$$|\partial_t \mathcal{E}(t, v)| \leq C_E(\mathcal{E}(t, v) + 1), \quad \forall t \in [0, T]. \quad (1.9)$$

Note that these assumptions closely mirror those of Theorem 1.3, which guarantees existence of solutions to the corresponding rate-independent system.

At this point, one can prove that any minimizer  $u_\lambda$  of  $\mathcal{I}_\lambda$  also satisfies the energy balance (E), that is

$$\mathcal{E}(t, u_\lambda(t)) + \int_{[0, t]} \Psi(\mathrm{d}u_\lambda) = \mathcal{E}(0, u_0) + \int_0^t \partial_s \mathcal{E}(s, u_\lambda(s)) \, \mathrm{d}s, \quad \forall t \in [0, T]. \quad (\text{E})$$

We observe that this identity for  $u_\lambda$  is entirely independent of  $\lambda$ ; as such, it can be used to obtain very useful *energetic a priori estimates* for the minimizers. Indeed, thanks to assumption (1.9) we have

$$\mathcal{E}(t, u_\lambda(t)) \leq (\mathcal{E}(0, u_0) + 1)e^{C_\mathcal{E}t} - 1$$

$$\int_{[0,t]} \Psi(du) \leq (\mathcal{E}(0, u_0) + 1)e^{C\mathcal{E}t}$$

From these estimates, and using the coercivity assumptions (1.7) and (1.8), one can obtain the following uniform a priori estimates

$$\exists C > 0, \forall \lambda > 0 : \quad \|u_\lambda\|_{L^\infty([0,T];X)} \leq C, \quad \int_{[0,T]} \|du\|_Z \leq C.$$

This suggests that the right choice to obtain precompactness of the minimizers could be

$$\mathcal{Y} = \text{BV}([0, T]; Z) \cap L^\infty([0, T]; X),$$

equipped with the weak\* convergence

$$u_k \xrightarrow{\mathcal{Y}} u \iff \begin{cases} \sup_k \int_{[0,T]} \|du_k\|_Z < \infty \\ \forall w \in L^1([0, T]; X^*) \\ \int_0^T \langle u_k(t), w(t) \rangle dt \rightarrow \int_0^T \langle u(t), w(t) \rangle dt \end{cases}$$

This is indeed the case; one can, in fact, prove the following convergence result for the minimizers, that will be the basis for our regularity result for rate-independent systems.

**Theorem 1.8** ([11], Theorem 4.2). *Any family of minimizers  $(u_\lambda)_{\lambda>0}$  for the family  $\mathcal{I}_\lambda$  is weakly\* precompact in  $\mathcal{Y}$ . Moreover, any limit point  $u \in \mathcal{Y}$  obtained for  $\lambda \rightarrow +\infty$  is a solution of the energetic formulation of the rate-independent system associated to  $\Psi$  and  $\mathcal{E}$ .*

After this first application to rate-independence processes, the same techniques have been applied to a wide range of evolution problems, ranging from gradient flows [14, 16] to Mean-Field games [2].



# Chapter 2

## Original contributions

As we anticipated in the Introduction the main contribution of this work is a Lipschitz regularity result for the minimizers of a class of weighted energy-dissipation functionals.

The possibility of uniform Lipschitz bounds for the minimizers of weighted energy dissipation functionals was suggested by a special case, studied in [6], which is inspired by a problem in the theory of Mean-Field games. In such an article, the authors derive uniform Lipschitz estimates for functionals of the form

$$F_\lambda(\rho) = \psi_0(\rho(0)) + \int_0^T e^{-\lambda t} \int_\Omega |\dot{\rho}(t, x)| + V(t, x)\rho(t, x) + f(\rho(t, x)) \, dx dt$$

among  $\rho \in \text{BV}([0, T], L^1(\Omega)) \cap L^2([0, T] \times \Omega)$  which are densities of probability measures on the set of unitary volume  $\Omega \subseteq \mathbb{R}^d$  (i.e. subject to the additional constraints  $\rho \geq 0$  and  $\int_\Omega \rho \, dx = 1, \forall t \geq 0$ ), where  $\psi_0$  is a given penalization for the initial value of the density, seeing that a Dirichlet condition of the type  $\rho(0) = \rho_0$ , is undesirable since  $\rho$ , being only of bounded variation, could have a jump at  $t = 0$ .

They prove that, under a uniform convexity assumption on  $f$ , i.e.  $f'' \geq c_0 > 0$ , a unique minimizer of  $F_\lambda$  exists and is Lipschitz continuous in time, independently of  $\lambda$ , with values in  $L^2(\Omega)$ .

Our main result generalizes this to the setting of abstract Banach spaces. In doing so, we extend the applicability of the result to the theory of rate-independent processes, providing new insights into the regularity of such systems.

In this very general context we were able to prove the following result.

**Theorem 2.1.** *Let  $(Z, \|\cdot\|_Z)$  be a real Banach space and  $X \subseteq Z$  be a subspace equipped with another norm  $\|\cdot\|_X$  such that:*

*$(X, \|\cdot\|_X)$  is a reflexive Banach space*

$\|\cdot\|_Z$  is lower semi-continuous with respect to the strong topology of  $X$   
 Let  $\mathcal{E} \in C^2([0, T] \times Z; [0, +\infty])$  and  $F : Z \rightarrow [0, +\infty]$  satisfy for all  $t \in [0, T]$ :

$\mathcal{E}(t, \cdot)$  is  $c_0$ -uniformly convex with respect to  $\|\cdot\|_X$

$\partial_t \mathcal{E}(t, \cdot)$  is  $c_0$ -Lipschitz continuous with respect to  $\|\cdot\|_X$

There exists  $C_\mathcal{E} > 0$  such that  $|\partial_t \mathcal{E}(t, x)| \leq C_\mathcal{E}(\mathcal{E}(t, x) + 1)$  for all  $x \in Z$

$F$  is convex and 1-Lipschitz-continuous with respect to  $\|\cdot\|_Z$

Finally, assume that there exists a family of finite dimensional subspaces  $(X_d)_{d \geq 1}$  of  $X$  such that for all  $x \in X$  one can find a sequence  $(x_d)_{d \geq 1}$  with  $x_d \in X_d$ ,  $x_d \rightarrow x$  with respect to  $\|\cdot\|_Z$ , and satisfying  $\mathcal{E}(t, x_d) \rightarrow \mathcal{E}(t, x)$  with  $\mathcal{E}(t, x_d) \leq \mathcal{E}(t, x) + 1$ , for all  $t \in [0, T]$ .

Then there exists a unique minimizer of

$$\mathcal{F}(v) = F(v(0^+)) + \int_{(0, T]} e^{-\lambda t} \|dv\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt$$

among curves  $v \in \text{BV}([0, T]; Z) \cap L^2([0, T]; X)$ . Moreover, the minimizer is Lipschitz continuous and satisfies

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0}, \quad \text{for a.e. } t \in [0, T]$$

Observe that we have not imposed an initial condition  $u(0) = u_0$  for the variational problem, mirroring the approach of [6]. This choice is motivated by the special role of Dirichlet boundary conditions in the BV setting: any curve  $u$  which takes the value  $u_0$  at  $t = 0$ , but has a different right-limit value  $u(0^+)$ , is still considered to satisfy the condition  $u(0) = u_0$ . In particular, one can freely choose the value of the curve on  $(0, T]$  to minimize the functional and then insert a jump to  $u_0$  at  $t = 0$ , adding a cost  $\|u(0^+) - u_0\|_Z$  to the total variation. Thus, we identify the value of the curve at  $t = 0$  with its right-limit and replace the Dirichlet condition with a penalization term  $F$ , where a typical example satisfying the theorem's assumptions is  $F(v) = \|v - u_0\|_Z$  for some  $u_0$ .

However, we also prove that if a Dirichlet initial condition is imposed, and the initial value satisfies an additional assumption

$$\|D\mathcal{E}(0, u_0)\|_{Z^*} \leq 1$$

then the minimizer has no jump at  $t = 0$ .

This condition means precisely  $u_0 \in \mathcal{S}(0)$ , the stable set at time  $t = 0$  of the rate-independent process associated to  $\Psi = \|\cdot\|_Z$  and  $\mathcal{E}$ . Indeed, as shown in Chapter 1, when the initial value is stable, we know that the solution to the rate-independent process is continuous at  $t = 0$ .

**Theorem 2.2.** *Consider the functional*

$$\mathcal{F}(v) = \int_{[0,T]} e^{-\lambda t} \|dv\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt$$

*among curves  $v \in \text{BV}([0, T]; Z) \cap L^2([0, T]; X)$  such that  $v(0) = u_0 \in X$ . Assume that, in addition to the hypotheses of Theorem 2.1, one also has*

$$\|D\mathcal{E}(0, u_0)\|_{Z^*} \leq 1$$

*Then, there exists a unique minimizer of  $\mathcal{F}$ . Moreover, this minimizer is Lipschitz continuous on  $[0, T]$  and satisfies*

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0}, \quad \text{for a.e. } t \in [0, T]$$

Using this regularity result and Theorem 1.8, one can immediately obtain the following:

**Theorem 2.3.** *Under the assumptions on  $X, Z$  and  $\mathcal{E}$  of Section 1.4 and those required in Theorem 2.1, there exists a unique energetic solution of the rate-independent system associated to  $\Psi = \|\cdot\|_Z$  and  $\mathcal{E}$  with initial value  $u_0 \in \mathcal{S}(0)$ . Moreover, the solution is Lipschitz continuous on  $[0, T]$  and satisfies*

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0}, \quad \text{for a.e. } t \in [0, T].$$

This can be seen as a generalization of Theorem 1.4, proved in Section 1.2, to the setting of Theorem 1.3. Notably, our result applies only to the case where  $\Psi$  is given by a norm  $\|\cdot\|_Z$ . However, we expect that, with some additional technical work, our proof could be extended to the more general setting discussed in Section 1.4.



# Chapter 3

## Proofs

### 3.1 Proof of Theorem 2.1

As we anticipated in the previous Chapter, the proof of Theorem 2.1 will consist of several approximation steps. In Section 3.1.1, we state and prove an analogue of Theorem 2.1 in finite-dimensional spaces, under the additional assumption that the norms  $\|\cdot\|_Z$  and  $\|\cdot\|_X$  are smooth. In Section 3.1.2, we present a way of approximating nonsmooth norms by smooth ones, which will allow us to extend our regularity result in finite dimension to generic nonsmooth norms  $\|\cdot\|_Z, \|\cdot\|_X$ . Finally, in Section 3.1.3, we prove Theorem 2.1 by using some finite dimensional approximations for which the regularity results of Section 3.1.1 apply.

#### 3.1.1 Regularity in finite dimension

In this section we prove the regularity result for a simpler problem on a finite dimensional real vector space  $V$  equipped with two smooth norms  $\|\cdot\|_Z, \|\cdot\|_X$ . Recall that, after having chosen a basis  $(e_i)_i$  of  $V$ , we can define a scalar product  $\langle \cdot, \cdot \rangle$  on  $V$  (and its associated norm  $|\cdot|$ ) by simply imposing  $\langle e_i, e_j \rangle = \delta_{ij}$  for  $i, j \in \{1, \dots, d\}$  ( $d$  being the dimension of  $V$  over  $\mathbb{R}$ ), with this structure  $V$  is isometric to  $\mathbb{R}^d$  when equipped with the euclidean scalar product. In what follows, the gradients (and higher order derivatives) of functions over  $V$  will be defined with respect to this linear structure; the same is true for the dual norms  $\|\cdot\|_{Z^*}, \|\cdot\|_{X^*}$ . For example for  $w \in V$  we have

$$\|w\|_{Z^*} = \sup\{\langle w, v \rangle : \|v\|_Z \leq 1\}.$$

We also recall that, since  $V$  is finite dimensional, any two norms  $\|\cdot\|_1, \|\cdot\|_2$  are equivalent, i.e. there exists a constant  $A \geq 1$  such that

$$\frac{1}{A}\|v\|_1 \leq \|v\|_2 \leq A\|v\|_1, \quad \forall v \in V,$$

and, thus, their respective topologies coincide. However, in some cases, we will still indicate the norm we are considering, by adding a subscript, as we did for infinite dimensional spaces; for instance we will write  $u \in L_X^2([0, T]; V)$  if

$$\int_0^T \|u(t)\|_X^2 dt < +\infty;$$

We are now ready to state and prove the main result of this Section.

**Theorem 3.1.** (*Regularity in finite dimensional spaces*) *Let  $V$  be finite dimensional real vector space, equipped with two norms  $\|\cdot\|_Z, \|\cdot\|_X$ , such that  $\|\cdot\|_X, \|\cdot\|_Z \in C^\infty(V \setminus \{0\})$  and  $\|\cdot\|_X^2$  is uniformly convex with respect to  $|\cdot|$ , i.e.  $\exists \gamma > 0$  such that  $\langle D^2(\|\cdot\|_X^2)(v)w, w \rangle \geq \gamma|w|^2$ , for all  $v, w \in V$ . Let  $\mathcal{E} \in C^2([0, T] \times V; [0, +\infty])$  be an energy-storage functional such that, for some constants  $c_0, C_0 > 0$ ,*

$$\langle D^2\mathcal{E}(t, v)w, w \rangle \geq c_0\|w\|_X^2, \quad \text{and} \quad \|\partial_t D\mathcal{E}(t, v)\|_{X^*} \leq C_0,$$

for all  $v, w \in V$  and  $t \in [0, T]$ . Let  $F : V \rightarrow [0, +\infty)$  be a 1-Lipschitz function with respect to  $\|\cdot\|_Z$ .

Then, there exists a unique minimizer  $u$  of

$$\mathcal{F}(v) = F(v(0^+)) + \int_{(0, T]} e^{-\lambda t} \|dv\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt,$$

over  $BV_Z([0, T], V) \cap L_X^2([0, T]; V)$ . Moreover, the minimizer is Lipschitz continuous and satisfies the estimate

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0}, \quad \text{for a.e. } t \in [0, T]. \quad (3.1)$$

*Proof.* We start by proving the desired regularity result for the regularized functionals  $\mathcal{F}_{\delta, \varepsilon, \alpha} : H^1([0, T]; V) \rightarrow [0, +\infty)$ , defined by

$$\mathcal{F}_{\delta, \varepsilon, \alpha}(v) = \alpha F(v(0^+)) + \int_0^T e^{-\lambda t} [L_{\delta, \varepsilon}(\dot{v}(t)) + \lambda \mathcal{E}(t, v(t))] dt,$$

where  $L_{\delta, \varepsilon} \in C^\infty(V \setminus \{0\}) \cap C^1(V)$  is an approximation of  $\|\cdot\|_Z$ , defined by

$$L_{\delta, \varepsilon}(v) = \sqrt{\delta^2 + \|v\|_Z^2} + \frac{\varepsilon}{2} \|v\|_X^2, \quad \forall v \in V.$$

The existence and uniqueness of a minimizer of  $\mathcal{F}_{\delta, \varepsilon, \alpha}$  can be readily established using the direct method. Indeed, one can obtain a uniform bound in  $H^1([0, T]; V)$

for any minimizing sequence, by leveraging the definition of  $L_{\delta,\varepsilon}$  and the uniform convexity of  $\mathcal{E}(t, \cdot)$ ; thus, we obtain precompactness of the minimizing sequence with respect to the weak topology of  $H^1([0, T]; V)$ . Finally, using the continuity and the convexity of  $L_{\delta,\varepsilon}$ ,  $\mathcal{E}(t, \cdot)$  and  $F$ , we can see that  $\mathcal{F}_{\delta,\varepsilon,\alpha}$  is sequentially lower semi-continuous with respect to the weak topology of  $H^1([0, T]; V)$ . The uniqueness of the minimizer follows because of the uniform convexity of  $\mathcal{E}(t, \cdot)$ .

In the following, when no confusion can arise, for fixed  $\delta, \varepsilon$  and  $\alpha$ , we will denote by  $u$  the unique minimizer of  $\mathcal{F}_{\delta,\varepsilon,\alpha}$ , instead of using the more precise notation  $u_{\delta,\varepsilon,\alpha}$ .

Let us now turn to the proof of the uniform bound (3.1). We begin by observing that the optimality conditions for  $u$  are given by the following system

$$\begin{cases} DL_{\delta,\varepsilon}(\dot{u}(t)) + D\mathcal{E}(t, u(t)) - \frac{1}{\lambda}D^2L_{\delta,\varepsilon}(\dot{u}(t))\ddot{u}(t) = 0, & \forall t \in (0, T), \\ DL_{\delta,\varepsilon}(\dot{u}(0)) = \alpha DF(u(0)), \\ DL_{\delta,\varepsilon}(\dot{u}(T)) = 0; \end{cases}$$

the first equation involves derivatives of  $u$  of order higher than one; this is not a problem, since using the regularity and uniform convexity of  $\mathcal{E}$  and  $L_{\delta,\varepsilon}$ , we obtain that  $u$  is differentiable up to three times (at points where  $\dot{u} \neq 0$ ). In particular, with our choice of  $L_{\delta,\varepsilon}$ , we see that  $\dot{u}$  is continuous in  $[0, T]$ .

Let us now introduce an auxiliary function  $h \in C^\infty(V \setminus \{0\}) \cap C^0(V)$ , defined by

$$h(v) = \langle DL_{\delta,\varepsilon}(v), v \rangle - L_{\delta,\varepsilon}(v),$$

and look for a point of maximum  $t_0 \in [0, T]$  of  $h(\dot{u}(t))$ , which exists because of the regularity of  $\dot{u}$  and  $h$ . We will prove that we have

$$\|\dot{u}(t_0)\|_X \leq C_{\delta,\varepsilon,\alpha},$$

for some constant  $C_{\delta,\varepsilon,\alpha} \geq 0$  such that

$$C_{\delta,\varepsilon,\alpha} \rightarrow \frac{C_0}{c_0}$$

for some appropriate choice of parameters  $\delta, \varepsilon \rightarrow 0^+$  and  $\alpha \rightarrow 1^-$ . From this, a direct computation gives

$$h(v) = -\frac{\delta^2}{\sqrt{\delta^2 + \|v\|_Z^2}} + \frac{\varepsilon}{2}\|v\|_X^2,$$

hence we obtain

$$h(\dot{u}(t_0)) \leq \frac{\varepsilon}{2}\|\dot{u}(t_0)\|_X^2 \leq \frac{\varepsilon}{2}C_{\delta,\varepsilon,\alpha}^2.$$

By definition of  $t_0$ , for any  $t \in [0, T]$  we have

$$-\frac{\delta^2}{\sqrt{\delta^2 + \|\dot{u}(t)\|_Z^2}} + \frac{\varepsilon}{2} \|\dot{u}(t)\|_X^2 = h(\dot{u}(t)) \leq h(\dot{u}(t_0)) \leq \frac{\varepsilon}{2} C_{\delta, \varepsilon, \alpha}^2,$$

which implies

$$\|\dot{u}(t)\|_X^2 \leq C_{\delta, \varepsilon, \alpha}^2 + \frac{2\delta}{\varepsilon}, \quad \forall t \in [0, T]. \quad (3.2)$$

At this point, observe that the right hand side of (3.2) is bounded provided one also impose

$$\frac{\delta}{\varepsilon} \rightarrow 0.$$

This will be used later in order to obtain equi-continuity for the family of minimizers.

We have two possible cases to analyze:  $t_0 \in (0, T)$  or  $t_0 \in \{0, T\}$ . We will treat each one separately.

Let us first suppose that  $t_0 \in (0, T)$ , we want to obtain a bound on  $\|\dot{u}(t_0)\|_X$ . If  $\dot{u}(t_0) = 0$  then we are done, otherwise, since

$$Dh(v) = D^2L_{\delta, \varepsilon}(v)v,$$

the optimality conditions for  $h(\dot{u}(t))$  at  $t = t_0$  read

$$\langle D^2L_{\delta, \varepsilon}(\dot{u})\dot{u}, \dot{u} \rangle = 0,$$

and

$$\langle D^2L_{\delta, \varepsilon}(\dot{u})\ddot{u}, \dot{u} \rangle + \langle D^3L_{\delta, \varepsilon}(\dot{u})[\ddot{u}, \dot{u}], \dot{u} \rangle + \langle D^2L_{\delta, \varepsilon}(\dot{u})\ddot{u}, \ddot{u} \rangle \leq 0.$$

(where  $\dot{u}$  and its derivatives are computed at  $t = t_0$ )

Let us now differentiate in time the Euler-Lagrange equation, compute it at  $t = t_0$  and take the duality product with  $\dot{u}(t_0) \in V$ . One then obtains (again omitting the evaluations at  $t = t_0$  for  $u$  and its derivatives)

$$\begin{aligned} 0 &= \langle D^2L_{\delta, \varepsilon}(\dot{u})\ddot{u}, \dot{u} \rangle + \langle D^2\mathcal{E}(t_0, u)\dot{u}, \dot{u} \rangle + \langle \partial_t D\mathcal{E}(t_0, u), \dot{u} \rangle \\ &\quad - \frac{1}{\lambda} \langle D^3L_{\delta, \varepsilon}(\dot{u})[\ddot{u}, \ddot{u}], \dot{u} \rangle - \frac{1}{\lambda} \langle D^2L_{\delta, \varepsilon}(\dot{u})\ddot{u}, \dot{u} \rangle \\ &\geq \langle D^2\mathcal{E}(t_0, u)\dot{u}, \dot{u} \rangle + \langle \partial_t D\mathcal{E}(t_0, u), \dot{u} \rangle + \frac{1}{\lambda} \langle D^2L_{\delta, \varepsilon}(\dot{u})\ddot{u}, \ddot{u} \rangle \\ &\geq -C_0\|\dot{u}\|_X + c_0\|\dot{u}\|_X^2 \end{aligned}$$

where one uses respectively, the optimality conditions at  $t = t_0$ , the assumptions on  $\mathcal{E}(t, \cdot)$  and the convexity of  $L_{\delta, \varepsilon}$ . So in conclusion one obtains

$$\|\dot{u}(t_0)\|_X \leq \frac{C_0}{c_0},$$

which is the desired estimate for  $\|\dot{u}(t_0)\|_X$ .

Let us now consider the case  $t_0 \in \{0, T\}$ . More specifically, we will only analyze the case  $t_0 = 0$  as the case  $t = T$  is analogous. If  $\dot{u}(0) = 0$  there is nothing to prove, otherwise we obtain the desired estimate on  $\|\dot{u}(0)\|_X$  in two steps. First, we prove the preliminary estimate

$$\|\dot{u}(0)\|_X \leq \frac{C\sqrt{\lambda}}{\sqrt{\varepsilon}}; \quad (3.3)$$

where  $C > 0$  is some constant independent of  $\varepsilon$ . Note that this is already an estimate on  $\|\dot{u}(0)\|_X$ , however the right hand side deteriorates as  $\varepsilon \rightarrow 0^+$ . Hence, in the second step, we will refine this estimate by using the transversality condition at  $t = 0$ .

In order to show that (3.3) holds, consider again the Euler-Lagrange equation on  $(0, T)$  and take the scalar product with  $\ddot{u}(t)$  to get

$$\langle D^2 L_{\delta, \varepsilon}(\dot{u}(t))\ddot{u}(t), \ddot{u}(t) \rangle = \lambda \langle D\mathcal{E}(t, u(t)) + DL_{\delta, \varepsilon}(\dot{u}(t)), \ddot{u}(t) \rangle.$$

At this point by using the uniform convexity assumptions on  $\|\cdot\|_X^2$  and  $\mathcal{E}(t, \cdot)$  we obtain

$$\langle D^2 L_{\delta, \varepsilon}(\dot{u}(t))\ddot{u}(t), \ddot{u}(t) \rangle \geq \frac{\varepsilon}{2} \langle D^2(\|\cdot\|_X^2)(\dot{u}(t))\ddot{u}(t), \ddot{u}(t) \rangle \geq \frac{C_1 \varepsilon \gamma}{2} \|\ddot{u}(t)\|_Z^2,$$

and

$$\begin{aligned} \langle D\mathcal{E}(t, u(t)) + DL_{\delta, \varepsilon}(\dot{u}(t)), \ddot{u}(t) \rangle &\leq \|\ddot{u}(t)\|_Z (\|D\mathcal{E}(t, u(t)) + DL_{\delta, \varepsilon}(\dot{u}(t))\|_{Z^*}) \\ &\leq C_2 \|\ddot{u}(t)\|_Z (1 + \|u(t)\|_Z) \\ &\leq C_2(1 + C_3) \|\ddot{u}(t)\|_Z, \end{aligned}$$

where  $C_1, C_2, C_3 > 0$  are some dimensional constants for which  $|\cdot|^2 \geq C_1 \|\cdot\|_Z^2$ ,  $\|D\mathcal{E}(t, u(t)) + DL_{\delta, \varepsilon}(\dot{u}(t))\|_{Z^*} \leq C_2(1 + \|u(t)\|_Z)$  and  $\|u(t)\|_Z \leq C_3$  (this constant exists since the minimality of  $u$  allows us to obtain a uniform bound on its total variation, thus in particular on  $\sup_t \|u(t)\|_Z$ ). In conclusion we have

$$\|\ddot{u}(t)\|_Z \leq \frac{c\lambda}{\varepsilon}, \quad \forall t \in (0, T).$$

for some constant  $c = \frac{2C_2(1+C_3)}{C_1\gamma} > 0$ .

Thus by Taylor's theorem we have

$$\|\dot{u}(t)\|_Z > \frac{1}{2} \|\dot{u}(0)\|_Z \quad \text{for } t < \frac{\varepsilon \|\dot{u}(0)\|_Z}{2c\lambda} = \hat{t},$$

hence, using the minimality of  $u$  and  $L_{\delta,\varepsilon}(v) \geq \|v\|_Z$ , we obtain

$$C_5 \geq \int_0^T L_{\delta,\varepsilon}(\dot{u}(t)) dt \geq \int_0^{\hat{t}} L_{\delta,\varepsilon}(\dot{u}(t)) dt \geq \int_0^{\hat{t}} \|\dot{u}(t)\|_Z dt > \frac{\varepsilon \|\dot{u}(0)\|_Z^2}{4c\lambda}$$

which finally implies the existence of a constant  $C > 0$  such that (3.3) holds. Let us now consider the transversality condition at  $t = 0$  for the minimizer  $u$ , which reads

$$DL_{\delta,\varepsilon}(\dot{u}(0)) = \alpha DF(u(0)).$$

We now use the fact that  $F$  is 1-Lipschitz with respect to  $\|\cdot\|_Z$ , which is equivalent to  $\|DF(v)\|_{Z^*} \leq 1$ , for all  $v \in V$ ; this implies

$$\|DL_{\delta,\varepsilon}(\dot{u}(0))\|_{Z^*} \leq \alpha < 1.$$

A direct computation gives

$$DL_{\delta,\varepsilon}(v) = \frac{\|v\|_Z D(\|\cdot\|_Z)(v)}{\sqrt{\delta^2 + \|v\|_Z^2}} + \varepsilon \|v\|_X D(\|\cdot\|_X)(v)$$

for  $v \neq 0$ ; which, in turn, implies

$$\begin{aligned} \frac{\|v\|_Z}{\sqrt{\delta^2 + \|v\|_Z^2}} - \|DL_{\delta,\varepsilon}(v)\|_{X^*} &\leq \left| \frac{\|v\|_Z \overbrace{\|D(\|\cdot\|_Z)(v)\|_{Z^*}}^{=1}}{\sqrt{\delta^2 + \|v\|_Z^2}} - \|DL_{\delta,\varepsilon}(v)\|_{X^*} \right| \\ &\leq \varepsilon \|v\|_X \|D(\|\cdot\|_X)(v)\|_{Z^*} \leq K\varepsilon \|v\|_X \end{aligned}$$

where we  $K > 0$  such that  $\frac{1}{K}\|\cdot\|_Z \leq \|\cdot\|_X \leq K\|\cdot\|_Z$  (hence also  $\frac{1}{K}\|\cdot\|_{X^*} \leq \|\cdot\|_{Z^*} \leq K\|\cdot\|_{X^*}$ ).

In conclusion we get

$$\frac{\|\dot{u}(0)\|_Z}{\sqrt{\delta^2 + \|\dot{u}(0)\|_Z^2}} \leq \alpha + K\varepsilon \|\dot{u}(0)\|_X \stackrel{(3.3)}{\leq} \alpha + KC\sqrt{\lambda\varepsilon}.$$

which implies

$$\|\dot{u}(0)\|_X^2 \leq K \frac{\delta^2(\alpha + KC\sqrt{\lambda\varepsilon})^2}{1 - (\alpha + KC\sqrt{\lambda\varepsilon})^2}; \quad (3.4)$$

observe now that, by choosing  $\delta$  and  $\alpha$  appropriately with respect to  $\varepsilon$ , the right hand side in the previous estimate can be made arbitrarily close to zero.

In conclusion we have been able to prove that for  $\varepsilon, \delta, \alpha$  fixed, the minimizer of  $\mathcal{F}_{\delta,\varepsilon,\alpha}$  satisfies

$$\|\dot{u}(t_0)\|_X^2 \leq C_{\delta,\varepsilon,\alpha}^2 = \max \left\{ \frac{C_0^2}{c_0^2}, K \frac{\delta^2(\alpha + KC\sqrt{\lambda\varepsilon})^2}{1 - (\alpha + KC\sqrt{\lambda\varepsilon})^2} \right\}. \quad (3.5)$$

To complete the proof of the Theorem we will show that with  $\delta_\varepsilon = \varepsilon^2$  and  $\alpha_\varepsilon = 1 - \varepsilon^{\frac{1}{4}}$ , the minimizers  $(u_\varepsilon)_{\varepsilon>0}$  of  $(\mathcal{F}_\varepsilon)_{\varepsilon>0}$  over  $H^1([0, T]; V)$  (that now depend only on the parameter  $\varepsilon$ ), converge uniformly to the unique minimizer  $u$  of  $\mathcal{F}$  over  $BV([0, T]; V) \cap L^2([0, T]; V)$ .

Indeed, as  $\varepsilon \rightarrow 0^+$ , we have  $\delta_\varepsilon \rightarrow 0^+$ ,  $\alpha_\varepsilon \rightarrow 1^-$  and

$$\frac{\delta_\varepsilon}{\varepsilon} \rightarrow 0 \quad \text{and} \quad K \frac{\delta_\varepsilon^2 (\alpha_\varepsilon + KC\sqrt{\lambda\varepsilon})^2}{1 - (\alpha_\varepsilon + KC\sqrt{\lambda\varepsilon})^2} \rightarrow 0$$

then thanks to (3.5) and (3.2), for  $\varepsilon$  small enough, we have

$$\|\dot{u}_\varepsilon(t)\|_X \leq \frac{C_0}{c_0} + 1, \quad \forall t \in [0, T]$$

this is saying that the family of minimizers  $(u_\varepsilon)_{\varepsilon>0}$  is equi-Lipschitz continuous. Moreover, by comparing  $u_\varepsilon$  with the constant curve  $\bar{u} = 0$  and using the uniform convexity of  $\mathcal{E}(t, \cdot)$  we can obtain the following uniform bound

$$\int_0^T \|u_\varepsilon(t)\|^2 dt \leq M,$$

where  $M$  is some positive constant independent of  $\varepsilon$ ; this, together with the equi-continuity just proved, implies equi-boundedness of  $(u_\varepsilon)_{\varepsilon>0}$ . So by Ascoli-Arzelà's theorem, there exists a function  $u \in C([0, T]; V)$  such that up to subsequences  $u_\varepsilon \rightrightarrows u$  in  $[0, T]$  as  $\varepsilon \rightarrow 0^+$  and  $\alpha \rightarrow 1^-$ . This curve is Lipschitz continuous and satisfies the desired estimate (3.1). We now prove that it is indeed the unique minimizer of  $\mathcal{F}$  over  $BV([0, T]; V) \cap L^2([0, T]; V)$ . The optimality of  $u_\varepsilon$ , together with  $L_\varepsilon(v) \geq \|v\|_Z$  implies that, for every curve  $w \in H^1([0, T]; V)$ , we have

$$\begin{aligned} \alpha_\varepsilon F(u_\varepsilon(0^+)) + \int_{(0, T]} e^{-\lambda t} \|du_\varepsilon\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u_\varepsilon(t)) dt \\ \leq \alpha_\varepsilon F(w(0^+)) + \int_0^T e^{-\lambda t} L_\varepsilon(\dot{w}(t)) dt + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, w(t)) dt. \end{aligned}$$

If we now send  $\varepsilon \rightarrow 0^+$  and use the lower semi-continuity of the weighted total variation integral with respect to uniform convergence, we obtain

$$\begin{aligned} F(u(0^+)) + \int_{(0, T]} e^{-\lambda t} \|du\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) dt \\ \leq F(w(0^+)) + \int_0^T e^{-\lambda t} \|dw\|_Z dt + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, w(t)) dt, \end{aligned}$$

where we have also used  $\alpha_\varepsilon \rightarrow 1$ ,  $u_\varepsilon(0^+) \rightarrow u(0^+)$ ,  $L_\varepsilon(v) \rightarrow \|v\|_Z$  and the dominated convergence theorem. This proves that  $u$  is a minimizer of  $\mathcal{F}$  over  $H^1([0, T]; V)$ . To conclude it is now enough to approximate any  $BV \cap L^2$   $w$  curve by  $H^1$  curves  $(w_k)_k$  such that

$$\limsup_k \mathcal{F}(w_k) \leq \mathcal{F}(w) \quad (3.6)$$

A simple way of obtaining  $w_k$  is by convolution of  $w$  with a smooth compactly supported kernel  $\rho_k$  tending to the identity. Indeed  $w_k = w * \rho_k$  is uniformly bounded and converges to  $w$  at all its continuity points, so in particular the convergence holds almost everywhere and at the boundary points. Moreover, we also have

$$\int_{(0, T]} e^{-\lambda t} \|dw_k\|_Z = \int_{(0, T]} e^{-\lambda t} \|d(w * \rho_k)\|_Z \leq \int_{(0, T]} e^{-\lambda t} \|dw\|_Z,$$

which, together with the dominated convergence theorem, allows us to obtain (3.6). We have just proved the existence of a minimizer of  $\mathcal{F}$  over  $BV([0, T]; V) \cap L^2([0, T]; V)$ , the uniqueness follows, as before, from the uniform convexity hypotheses on  $\mathcal{E}(t, \cdot)$ .  $\square$

### 3.1.2 Approximation of norms

One major requirement in the proof of Theorem 3.1 is the smoothness and uniform convexity of the norms involved. Clearly, two generic norms  $\|\cdot\|_Z$  and  $\|\cdot\|_X$  over a real vector space do not satisfy this assumption. Consider, for instance, the  $\ell^1$  norm over  $\mathbb{R}^2$ , which is not differentiable on the axes  $\{x_i = 0\}$ , or the  $\|\cdot\|_{\ell^p}$  norm over  $\mathbb{R}^2$  when  $p > 2$ , for which its square is not uniformly convex.

This complication can be resolved by approximating the norms of our problem by smooth and uniformly convex norms, in such a way that the functional  $\mathcal{F}$  computed with these new norms  $\Gamma$ -converges to the one with the original norms in some appropriate topology. This will be the content of Proposition 3.1.1.

However, before giving a proof of the  $\Gamma$ -convergence result, let us present a possible way of performing the approximation of the norms.

In what follows  $\|\cdot\|$  denotes a general norm on a finite dimensional real normed vector space  $X$ . Let us recall that, after having fixed a basis  $(e_i)_{1 \leq i \leq d}$ , we can easily define a scalar product  $\langle \cdot, \cdot \rangle$  on  $X$  which makes  $X$  isometric to  $\mathbb{R}^d$  ( $d$  being the dimension of  $X$ ). Thanks to this isometry one can also define a measure over  $X$ , such that the unit cube  $\{\sum_{i=1}^d x_i e_i : |x_i| \leq 1, \forall i\}$  has unit measure. As before, we will denote by  $|\cdot|$  the euclidean norm associated to this scalar product on  $X$ , while the ball of center 0 and radius  $r$  with respect to  $|\cdot|$  will be denote by  $B_r$ .

What we will show is that for every  $\sigma > 0$  small enough, there exists a norm  $\|\cdot\|_\sigma$

such that for every vector  $v$  it satisfies

$$\|v\|_\sigma = (1 + o_\sigma(1))\|v\|, \quad \text{for } \sigma \rightarrow 0^+. \quad (3.7)$$

We will construct the norms  $\|\cdot\|_\sigma$  by approximating the unit ball  $\mathcal{B} = \{x \in X : \|x\| \leq 1\}$  of the original norm  $\|\cdot\|$  by smooth and uniformly convex sets. These sets will be constructed as sublevel sets of smooth and strongly log-concave functions that approximate the indicator function  $\mathbf{1}_\mathcal{B}$  of  $\mathcal{B}$ , defined by

$$\mathbf{1}_\mathcal{B}(x) = \begin{cases} 1 & \text{for } x \in \mathcal{B}, \\ 0 & \text{for } x \notin \mathcal{B}. \end{cases}$$

A common way of obtaining smooth approximations of a function is by convolution with a smooth kernel. Let us define  $\mathcal{N}_\sigma : X \rightarrow [0, +\infty)$ , the density of the normal distribution over  $X$  of variance  $\sigma^2$ , by

$$\mathcal{N}_\sigma(x) = \frac{1}{(2\pi\sigma)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{2\sigma}\right);$$

in particular, we have  $\mathcal{N}_\sigma \in C^\infty(X)$ ,  $\mathcal{N}_\sigma(x) > 0$ ,  $\mathcal{N}_\sigma(-x) = \mathcal{N}_\sigma(x)$  for all  $x \in X$  and  $\int_X \mathcal{N}_\sigma(y) dy = 1$ . Then we define  $U_\sigma : X \rightarrow [0, +\infty)$  by

$$U_\sigma(x) = \left[ (e^{-\sigma|\cdot|^2} \mathbf{1}_\mathcal{B}) * \mathcal{N}_\sigma \right] (x) = \int_X e^{-\sigma|x-y|^2} \mathbf{1}_\mathcal{B}(x-y) \mathcal{N}_\sigma(y) dy.$$

This will be our smooth approximation of the indicator function  $\mathbf{1}_\mathcal{B}$ . Indeed, by properties of convolution  $U_\sigma \in C^\infty(X)$ ,  $U_\sigma(x) > 0$ ,  $U_\sigma(-x) = U_\sigma(x)$  for all  $x \in X$ . Moreover  $U_\sigma = \exp(-V_\sigma)$  for some uniformly convex function  $V_\sigma \in C^\infty(X)$ ; this is because  $U_\sigma$  is the convolution of two functions that have this property, which in the literature are usually called strongly log-concave functions (for a proof of this see for instance [17]); as such it possesses a unique maximum point, which by symmetry has to be 0. We now have the following Lemma:

**Lemma 3.1.1.** *Let  $V \in C^\infty(X)$  be such that  $V(x) > 0$ ,  $V(x) = V(-x)$  and  $D^2V(x) \geq \alpha I$  for some  $\alpha > 0$ , for all  $x \in X$ . For some  $c > 0$ , define  $f : X \rightarrow [0, +\infty)$  by*

$$f(x) = \inf\{t > 0 : V(tx) < c\}, \quad \forall x \in X,$$

*then  $f$  is a norm on  $X$  such that  $f \in C^\infty(X \setminus \{0\})$  and  $f^2$  is uniformly convex with respect to  $|\cdot|$ .*

*Proof.* We only prove the uniform convexity of  $f^2$ . Directly from the definition we get that for all  $x \in X \setminus \{0\}$  we have

$$V\left(\frac{x}{f(x)}\right) = c;$$

by taking the gradient of the previous equation we obtain

$$DV = \frac{\langle DV\left(\frac{x}{f}\right), x \rangle}{f} Df; \quad (3.8)$$

in the previous identity and for the subsequent ones, the function  $f$  and its derivatives are computed at  $x$ . After taking another gradient we obtain the following identity

$$\begin{aligned} & D^2V\left(\frac{x}{f}\right) - \frac{1}{f} \left( D^2V\left(\frac{x}{f}\right) x \right) \otimes Df = \\ &= \frac{1}{f} Df \otimes \left( D^2V\left(\frac{x}{f}\right) x \right) - \frac{1}{f} \left\langle D^2V\left(\frac{x}{f}\right) x, x \right\rangle Df \otimes Df \\ & \quad + Df \otimes DV\left(\frac{x}{f}\right) + \left\langle DV\left(\frac{x}{f}\right), x \right\rangle D^2f - \left\langle DV\left(\frac{x}{f}\right), x \right\rangle Df \otimes Df; \end{aligned}$$

if we use (3.8) and  $\frac{D^2(f^2)}{2} = fD^2f + Df \otimes Df$ , we conclude that

$$\begin{aligned} \left\langle DV\left(\frac{x}{f}\right), x \right\rangle \frac{D^2(f^2)}{2} &= D^2V\left(\frac{x}{f}\right) + \left\langle D^2V\left(\frac{x}{f}\right) x, x \right\rangle Df \otimes Df \\ & \quad - \left[ \left( D^2V\left(\frac{x}{f}\right) x \right) \otimes Df + Df \otimes \left( D^2V\left(\frac{x}{f}\right) x \right) \right] \\ & \quad + f \left\langle DV\left(\frac{x}{f}\right), x \right\rangle Df \otimes Df. \end{aligned}$$

To conclude we just have to observe that by construction

$$\left\langle DV\left(\frac{x}{f}\right), x \right\rangle > 0,$$

and that for any vector  $w \in X$ , applying these bilinear terms to  $(w, w) \in X \times X$ , we have that the right hand side is equal to

$$f \left\langle DV\left(\frac{x}{f}\right), x \right\rangle ((Df, w))^2 + \left\langle D^2V\left(\frac{x}{f}\right) [w - \langle Df, w \rangle x], [w - \langle Df, w \rangle x] \right\rangle$$

which is greater than  $\beta|w|^2$  for some  $\beta > 0$  using the uniform convexity of  $V$ .  $\square$

By applying Lemma 3.1.1 to  $V_\sigma$  with  $c = \log 2$ , we can define a smooth norm  $\|\cdot\|_\sigma$  over  $X$  such that its square is uniformly convex and for which the unit ball is given by the set

$$\{x \in X : V_\sigma(x) \leq \log 2\} = \left\{ x \in X : U_\sigma(x) \geq \frac{1}{2} \right\}$$

To conclude we just need to show that with this construction, (3.7) holds.

In the following  $C > 0$  denotes a constant for which  $\frac{1}{C}|x| \leq \|x\| \leq C|x|$  for all  $x \in X$  (such  $C$  exists since  $X$  is finite dimensional).

First, we observe that (3.7) is equivalent to proving that there exists  $\delta = o_\sigma(1)$  as  $\sigma \rightarrow 0^+$ , such that for  $\sigma$  small enough, the boundary of the unit ball of  $\|\cdot\|_\sigma$  lies in a  $\delta$ -neighborhood of the boundary of  $\mathcal{B}$ .

We will show that  $\delta = 2C\sigma$  works. Indeed, choose  $x \in X$  such that  $\|x\| \leq 1 - 2C\sigma$ , then we have

$$\begin{aligned} U_\sigma(x) &= \int_X e^{-\sigma|x-y|^2} \mathbf{1}_{\mathcal{B}}(x-y) \mathcal{N}_\sigma(y) \, dy \\ &\geq e^{-C^2\sigma} \int_{\{\|x-y\| \leq 1\}} \mathcal{N}_\sigma(y) \, dy \\ &\geq e^{-C^2\sigma} \int_{\{|y| \leq 2\sigma\}} \mathcal{N}_\sigma(y) \, dy \geq e^{-C^2\sigma} \frac{3}{4} > \frac{1}{2}, \end{aligned}$$

for  $\sigma$  small enough; similarly, choose  $x \in X$  such that  $\|x\| \geq 1 + 2C\sigma$ , then

$$\begin{aligned} U_\sigma(x) &= \int_X e^{-\sigma|x-y|^2} \mathbf{1}_{\mathcal{B}}(x-y) \mathcal{N}_\sigma(y) \, dy \\ &\leq \int_{\{\|x-y\| \leq 1\}} \mathcal{N}_\sigma(y) \, dy \\ &\leq \int_{\{|y| \geq 2\sigma\}} \mathcal{N}_\sigma(y) \, dy \leq \frac{1}{4} < \frac{1}{2}. \end{aligned}$$

Where for the previous estimates we have used that, for all  $k \in \mathbb{N}$ ,

$$\int_{\{|y| \geq 2k\sigma\}} \mathcal{N}_\sigma(y) \, dy \leq \frac{1}{k^2}.$$

We are now ready to state and prove the  $\Gamma$ -convergence result mentioned above. For all  $\sigma > 0$  small enough define the norms  $\|\cdot\|_{Z,\sigma}$  and  $\|\cdot\|_{X,\sigma}$ , using the construction just describe. Moreover, define

$$\mathcal{F}_\sigma(u) = F(u(0^+)) + \int_{(0,T]} e^{-\lambda t} \|du\|_{Z,\sigma} + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) \, dt$$

**Proposition 3.1.1.** *We have  $\mathcal{F}_\sigma \xrightarrow{\Gamma} \mathcal{F}$ , with respect to the uniform convergence in  $X$ .*

Recall that the notion of uniform convergence in  $X$  is independent of the norm, since at this point  $X$  is finite-dimensional. Secondly, observe that the family of

minimizers  $(u_\sigma)_\sigma$  of  $(\mathcal{F}_\sigma)_\sigma$  (over  $\text{BV}([0, T]; X) \cap \text{L}^2([0, T]; X)$ ) is precompact with respect to the uniform convergence. Indeed, with the previous construction we have that, for all  $x, w \in X, t \in [0, T]$ ,

$$\langle \text{D}^2 \mathcal{E}(t, x)w, w \rangle \geq \frac{c_0}{(1 + o_\sigma(1))^2} \|w\|_{X, \sigma}^2 \quad \text{and} \quad \|\partial_t \nabla \mathcal{E}(t, x)\|_{X', \sigma} \leq (1 + o_\sigma(1))C_0.$$

Thus, from Theorem 3.1, the minimizer  $u_\sigma$  of  $\mathcal{F}_\sigma$  satisfies

$$\|\dot{u}_\sigma(t)\|_{X, \sigma} \leq (1 + o_\sigma(1))^3 \frac{C_0}{c_0}, \quad \forall t \in [0, T].$$

So for  $\sigma$  small enough, using again  $\|\cdot\|_{X, \sigma} = (1 + o_\sigma(1))\|\cdot\|_X$ , we get

$$\|\dot{u}_\sigma(t)\|_X \leq \frac{C_0}{c_0} + 1, \quad \forall t \in [0, T],$$

uniformly in  $\sigma$ . Moreover, similarly to what was done in Theorem 3.1, by comparing  $u_\sigma$  with a constant curve, we can obtain a uniform bound of the form

$$\int_0^T \|u_\sigma(t)\|_X^2 dt \leq C,$$

which, together with the equicontinuity of  $u_\sigma$  with respect to  $\|\cdot\|_X$ , implies equiboundedness. Hence, we obtain the desired precompactness by Ascoli-Arzelà's theorem.

*Proof of Proposition 3.1.1.* The proof of the  $\Gamma$  – lim inf part of the Proposition is almost immediate. Indeed, if  $v_\sigma \rightrightarrows v$ , then  $v_\sigma(0^+) \rightarrow v(0^+)$ , which, together with the continuity of  $F$ , implies  $F(v_\sigma(0^+)) \rightarrow F(v(0^+))$ . Then, using  $\|\cdot\|_{Z, \sigma} = (1 + o_\sigma(1))\|\cdot\|_Z$  and the lower semi-continuity of the total variation with respect to the uniform convergence, we get

$$\int_{(0, T]} e^{-\lambda t} \|dv\|_Z \leq \liminf_\sigma \int_{(0, T]} e^{-\lambda t} \|dv_\sigma\|_{Z, \sigma}.$$

Finally, by Fatou's Lemma we also get

$$\int_{(0, T]} e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt \leq \liminf_\sigma \int_{(0, T]} e^{-\lambda t} \lambda \mathcal{E}(t, v_\sigma(t)) dt.$$

For the  $\Gamma$  – lim sup part of the  $\Gamma$ -limit we can just choose the constant recovery sequence. Indeed, given  $u : [0, T] \rightarrow Z$  such that  $\mathcal{F}(u) < +\infty$  (if infinite there is nothing to prove), define  $u_\sigma(t) = u(t)$  for all  $t \in [0, T]$ . Then, trivially  $u_\sigma \rightrightarrows u$ , and, using  $\|\cdot\|_{Z, \sigma} = (1 + o_\sigma(1))\|\cdot\|_Z$ , we also have

$$\limsup_\sigma \mathcal{F}_\sigma(u_\sigma) = \limsup_\sigma \mathcal{F}_\sigma(u) \leq \mathcal{F}(u),$$

which concludes the proof.  $\square$

As a consequence of Proposition 3.1.1, the precompactness of  $(u_\sigma)_\sigma$  with respect to the uniform convergence, and the estimate

$$\|\dot{u}_\sigma(t)\|_{X,\sigma} \leq (1 + o_\sigma(1))^2 \frac{C_0}{c_0},$$

we obtain the desired Lipschitz regularity result in finite dimension

**Theorem 3.2** (Lipschitz regularity in finite dimension). *Under the hypotheses on  $\mathcal{E}$  of Theorem 3.1, there exists a unique minimizer  $u$  of  $\mathcal{F}$  over  $\text{BV}([0, T]; Z) \cap L^2([0, T], X)$ . Moreover,  $u$  is Lipschitz continuous with respect to  $\|\cdot\|_X$  and satisfies the estimate*

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0},$$

for a.e.  $t \in [0, T]$ .

### 3.1.3 Regularity in infinite dimension

In this section we finally prove Theorem 2.1. As we already explained, the proof is based on the regularity result in finite dimension contained in Theorem 3.1 and a  $\Gamma$  convergence result for the functionals in finite dimension, contained in the following proposition.

**Proposition 3.2.1.** *Let  $(Z, \|\cdot\|_Z)$  be a real Banach space and  $X \subseteq Z$  be a subspace equipped with another norm  $\|\cdot\|_X$  such that*

- $(X, \|\cdot\|_X)$  is a reflexive Banach space;
- $\|\cdot\|_Z$  is lower semi-continuous with respect to the strong topology of  $X$ .

Let  $\mathcal{E} \in C^2([0, T] \times Z; [0, +\infty))$  and  $F : Z \rightarrow [0, +\infty)$  satisfy

- $\mathcal{E}(t, \cdot)$  is  $c_0$ -uniformly convex with respect to  $\|\cdot\|_X$ ;
- $\partial_t \mathcal{E}(t, \cdot)$  is  $C_0$ -Lipschitz continuous with respect to  $\|\cdot\|_X$ ;
- $|\partial_t \mathcal{E}(t, x)| \leq C_\mathcal{E}(\mathcal{E}(t, x) + 1)$
- $F$  is convex and 1-Lipschitz continuous with respect to  $\|\cdot\|_Z$ .

Finally, assume that

- There exists a family of finite dimensional subspaces  $(X_d)_{d \geq 1}$  of  $X$  such that for all  $x \in X$  one can find a sequence  $(x_d)_{d \geq 1}$  with  $x_d \in X_d$ ,  $x_d \rightarrow x$  with respect to  $\|\cdot\|_Z$ , and  $\mathcal{E}(t, x_d) \rightarrow \mathcal{E}(t, x)$  with  $\mathcal{E}(t, x_d) \leq \mathcal{E}(t, x) + 1$  for all  $t \in [0, T]$ .

Then  $\mathcal{F}_d \xrightarrow{\Gamma} \mathcal{F}$  with respect to the weak  $L^2_X$  convergence, where

$$\mathcal{F}_d(u) = F(u(0^+)) + \int_{(0,T]} e^{-\lambda t} \|du\|_Z + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) dt + I_{X_d}(u)$$

and

$$I_{X_d}(u) = \begin{cases} 0 & \text{if } u(t) \in X_d, \text{ for a.e. } t \in [0, T], \\ +\infty & \text{otherwise.} \end{cases}$$

*Proof of Proposition 3.2.1.* We begin with the proof of the  $\Gamma$ -liminf part of the  $\Gamma$ -limit. Let us first observe that it suffices to prove that  $\mathcal{F}$  is lower semi-continuous on  $\text{BV}([0, T]; Z) \cap L^2([0, T]; X)$  with respect to the weak  $L^2_X$  convergence; indeed, if this was true then for every  $v_d \rightharpoonup v$  in  $L^2_X$  we would have

$$\mathcal{F}(v) \leq \liminf_d \mathcal{F}(v_d) \leq \liminf_d \mathcal{F}_d(v_d),$$

since  $\mathcal{F} \leq \mathcal{F}_d$  for all  $d \geq 1$ .

In order to prove the lower semi-continuity of  $\mathcal{F}$ , let us fix a small parameter  $\varepsilon > 0$ , and consider the new functional

$$\mathcal{F}_\varepsilon(u) = e^{-\lambda\varepsilon} F\left(\int_0^\varepsilon u(t) dt\right) + V_\varepsilon(u; [\varepsilon, T]) + \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, u(t)) dt,$$

where

$$V_\varepsilon(u; [\varepsilon, T]) = \sup_{\substack{t_0=\varepsilon, t_N=T \\ t_{k+1} > t_k + \varepsilon}} \left\{ \sum_{k=1}^N e^{-t_k + \frac{\varepsilon}{2}} \|\bar{u}_{t_k} - \bar{u}_{t_{k-1}}\|_Z : \bar{u}_t = \int_{t-\frac{\varepsilon}{2}}^{t+\frac{\varepsilon}{2}} u(s) ds \right\}$$

We only need to prove lower semi-continuity of this new functional and later take the limit as  $\varepsilon \rightarrow 0^+$ . Since we have by Jensen's inequality

$$\begin{aligned} F\left(\int_0^\varepsilon u(t) dt\right) &\leq \int_0^\varepsilon F(u(t)) dt \leq F(u(0^+)) + \int_0^\varepsilon \|u(t) - u(0^+)\|_Z dt \\ &\leq F(u(0^+)) + \int_{(0,\varepsilon)} \|du\|_Z; \end{aligned}$$

hence

$$\begin{aligned} e^{-\lambda\varepsilon} F\left(\int_0^\varepsilon u(t) dt\right) &\leq e^{-\lambda\varepsilon} F(u(0^+)) + e^{-\lambda\varepsilon} \int_{(0,\varepsilon)} \|du\|_Z \\ &\leq F(u(0^+)) + \int_{(0,\varepsilon)} e^{-\lambda t} \|du\|_Z. \end{aligned}$$

And similarly

$$V_\varepsilon(u; [\varepsilon, T]) \leq \int_{(0, T]} e^{-\lambda t} \|du\|_Z$$

But the lower semi-continuity of  $\mathcal{F}_\varepsilon$  is immediate since if  $v_d \rightharpoonup v$  in  $L^2_X$  then, for all  $0 \leq s < t \leq T$ , one has

$$\int_s^t v_d(r) dr \rightarrow \int_s^t v(r) dr.$$

For the  $\Gamma$ -limsup part of the  $\Gamma$ -limit we would have to check that for any  $v \in \text{BV}([0, T]; Z) \cap L^2([0, T]; X)$  with  $\mathcal{F}(v) < +\infty$  there exists a sequence  $(v_d)_{d \geq 1}$  such that  $v_d \rightharpoonup v$  in  $L^2_X$  and

$$\limsup_d \mathcal{F}_d(v_d) \leq \mathcal{F}(v).$$

Let us assume, for now, that one just needs to check this for any  $v : [0, T] \rightarrow Z$  which is right-continuous, piecewise constant with a finite number of jumps and with  $\mathcal{F}(v) < +\infty$ . We recall that this last condition implies that  $v$  takes values in  $X$ .

In particular, we have

$$v(t) = v_k \in X, \quad \text{for } t \in [t_{k-1}, t_k), \quad 0 = t_0 < \dots t_N = T.$$

We have to find  $(v_d)_{d \geq 1}$  such that  $v_d \rightharpoonup v$  in  $L^2_X$  and

$$\limsup_d \mathcal{F}_d(v_d) \leq \mathcal{F}(v).$$

By the last assumption, for all  $k = 1, \dots, N$ , there exists a sequence  $(v_{k,d})_{d \geq 1}$  such that  $v_{k,d} \in X_d$ ,  $\|v_{k,d} - v_k\|_Z \rightarrow 0$ , and for which  $\mathcal{E}(t, v_{k,d}) \rightarrow \mathcal{E}(t, v_k)$  with  $\mathcal{E}(t, v_{k,d}) \leq \mathcal{E}(t, v_k) + 1$  for all  $t \in [0, T]$ . Define, then,  $v_d$  by

$$v_d(t) = v_{k,d} \in X_d, \quad \text{for } t \in [t_{k-1}, t_k), \quad 0 = t_0 < \dots t_N = T.$$

Clearly we have  $F(v_d(0^+)) \rightarrow F(v(0^+))$ , since  $F$  is 1-Lipschitz continuous with respect to  $\|\cdot\|_Z$ . Moreover, by working on each interval  $[t_{k-1}, t_k)$  separately, by strong convergence in  $Z$  we have

$$\limsup_d \int_{(0, T]} e^{-\lambda t} \|dv_d\|_Z \leq \int_{(0, T]} e^{-\lambda t} \|dv\|_Z,$$

and by dominated convergence we have

$$\lim_d \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v_d(t)) dt = \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt.$$

In conclusion, since with our choice  $I_{X_d}(v_d) = 0$  for all  $d \geq 1$ , we have proved

$$\limsup_d \mathcal{F}_d(v_d) \leq \limsup_d \mathcal{F}(v_d) \leq \mathcal{F}(v).$$

Now it only remains to prove  $v_d \rightharpoonup v$  in  $L^2_X$ . We clearly have  $v_d \rightarrow v$  in  $L^\infty_Z$ ; moreover, the uniform convexity of  $\mathcal{E}(t, \cdot)$  with respect to  $\|\cdot\|_X$  and the upper bound on  $\mathcal{F}(v_d)$  we just proved imply that  $(v_d)_{d \geq 1}$  is equi-bounded in  $L^2_X$ , so because of reflexivity of  $(X, \|\cdot\|_X)$  (up to taking a subsequence) there exists  $w \in L^2_X([0, T]; X)$  such that  $v_d \rightharpoonup w$  in  $L^2_X$ . If we are able to prove that  $v = w$  we are done.

To do so, it suffices to show that if family of vectors  $(x_d)_d$  converges strongly to  $x_1$  in  $Z$ , and weakly to  $x_2$  in  $X$ , then  $x_1 = x_2$ . Indeed, if this is true, then for any  $0 \leq s < t < T$ , the vectors

$$\int_s^t v_d(r) dr,$$

would converge strongly to  $\int_s^t v(r) dr$  in  $Z$  and weakly to  $\int_s^t w(r) dr$  in  $X$ . Thus

$$\int_s^t v(r) dr = \int_s^t w(r) dr,$$

which would imply  $v = w$  on  $[0, T]$  due to the arbitrariness of  $s$  and  $t$ . In order to prove the equality between the strong limit in  $Z$  and the weak limit in  $X$ , we can show that there exists a norm on  $X$  which is weaker than both  $\|\cdot\|_Z$  and  $\|\cdot\|_X$ . Consider the following definition:

$$\|x\|_{Z \oplus X} = \inf\{\|x_1\|_Z + \|x_2\|_X : x_1 + x_2 = x\}.$$

It is clear that  $\|\cdot\|_{Z \oplus X} \leq \|\cdot\|_Z, \|\cdot\|_X$ . Moreover,  $\|\cdot\|_{Z \oplus X}$  can be easily shown to define a norm on  $X$ . The only non-trivial property to verify is the positive definiteness of  $\|\cdot\|_{Z \oplus X}$ . Suppose a vector  $x \in X$  satisfies  $\|x\|_{Z \oplus X} = 0$ . Then, there exist sequences  $(x_n^1)_n$  and  $(x_n^2)_n$  such that  $x_n^1 + x_n^2 = x$  and

$$\|x_n^1\|_Z + \|x_n^2\|_X \rightarrow \|x\|_{Z \oplus X} = 0.$$

In particular  $x_n^1 \rightarrow 0$  in  $Z$ , and  $x_n^2 \rightarrow 0$  in  $X$ . It then follows immediately that  $x_n^1 \rightarrow x$  in  $X$ , since  $\|x - x_n^1\|_X = \|x_n^2\|_X \rightarrow 0$ . Hence, using the lower semi-continuity of  $\|\cdot\|_Z$  with respect to the strong topology of  $X$ , we conclude

$$\|x\|_Z \leq \liminf_n \|x_n^1\|_Z = 0,$$

which implies  $x = 0$ .

The reduction to right continuous piecewise constant curves with finite number of jumps is possible because, as we are about to show, the class  $\mathcal{P}$  of such curves is

energy dense in  $\text{BV}([0, T]; Z) \cap L^2([0, T]; X)$  for  $\mathcal{F}$  and the weak  $L^2_X$  topology; that is, for every  $v \in \text{BV}([0, T]; Z) \cap L^2([0, T]; X)$ , there exists a sequence  $(v_k)_k \subset \mathcal{P}$  such that  $v_k \rightharpoonup v$  in  $L^2_X$  and

$$\limsup_k \mathcal{F}(v_k) \leq \mathcal{F}(v).$$

(see [3] for an abstract presentation of this energy-density argument) Let  $v$  be a generic curve in  $\text{BV}_Z([0, T]; Z) \cap L^2([0, T]; X)$  with  $\mathcal{F}(v) < +\infty$  (otherwise there is nothing to prove) and fix  $\delta > 0$  small. Consider a partition  $0 = t_0 < \dots, t_N = T$  of  $[0, T]$  such that  $|t_k - t_{k-1}| \leq \delta$  for all  $k = 1, \dots, N$ . For each subinterval  $(t_{k-1}, t_k] = I_k$  we can choose a point  $\hat{t}_k$  such that

$$\mathcal{E}(\hat{t}_k, v(\hat{t}_k)) \leq \frac{\int_{I_k} e^{-\lambda t} \mathcal{E}(\hat{t}, v(\hat{t})) dt}{\int_{I_k} e^{-\lambda t} dt}.$$

Define  $v_\delta : [0, T] \rightarrow X$  by

$$v_\delta(t) = v(\hat{t}_k) \in X, \quad \text{for } t \in I_k, \quad k = 1, \dots, N.$$

We will prove that  $v_\delta \rightharpoonup v$  in  $L^2_X$  and

$$\limsup_{\delta \rightarrow 0^+} \mathcal{F}(v_\delta) \leq \mathcal{F}(v).$$

The argument to prove weak  $L^2_X$  convergence of  $v_\delta$  to  $v$  is analogous to the one we used in the  $\Gamma$ -lim sup above, with the minor difference that in this case we only have  $v_\delta \rightarrow v$  in  $L^1([0, T]; Z)$ , since for every  $k$  we have

$$\int_{I_k} \|v_\delta(t) - v(t)\|_Z dt = \int_{I_k} \|v(\hat{t}_k) - v(t)\|_Z dt \leq \delta \int_{I_k} \|dv\|_Z$$

and thus

$$\int_0^T \|v_\delta(t) - v(t)\|_Z dt \leq \delta e^{-\lambda T} \mathcal{F}(v) \rightarrow 0.$$

Moreover, we clearly have  $F(v_\delta(0^+)) \rightarrow F(v(0^+))$  since  $F$  is 1-Lipschitz continuous with respect to  $\|\cdot\|_Z$  and

$$\|v_\delta(0^+) - v(0^+)\|_Z \leq \int_{(0, \delta)} \|dv\|_Z \rightarrow 0.$$

Then, we also immediately have

$$\int_{(0, T]} e^{-\lambda t} \|dv_\delta\|_Z \leq e^{\lambda \delta} \int_{(0, T]} e^{-\lambda t} \|dv\|_Z.$$

It remains to prove that

$$\limsup_{\delta} \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v_{\delta}(t)) dt \leq \int_0^T e^{-\lambda t} \lambda \mathcal{E}(t, v(t)) dt \quad (3.9)$$

Let us first work on each interval  $I_k$  separately. With our choice of  $\hat{t}_k$ , we have that

$$\begin{aligned} \int_{I_k} e^{-\lambda t} \mathcal{E}(t, v_{\delta}(t)) dt &\leq \int_{I_k} e^{-\lambda t} \mathcal{E}(t, v(\hat{t}_k)) dt + \int_{I_k} e^{-\lambda t} \varphi(t) dt \\ &\leq \int_{I_k} e^{-\lambda t} \mathcal{E}(t, v(t)) dt + \int_{I_k} e^{-\lambda t} \varphi(t) dt, \end{aligned}$$

where  $\varphi(t) = |\mathcal{E}(t, v(\hat{t}_k)) - \mathcal{E}(\hat{t}_k, v(\hat{t}_k))|$  for  $t \in I_k$ . But then  $\varphi(\hat{t}_k) = 0$  and

$$\begin{aligned} |\varphi'(t)| &\leq |\partial_t \mathcal{E}(t, v(\hat{t}_k))| \leq C_{\mathcal{E}}(\mathcal{E}(t, v(\hat{t}_k)) + 1) \\ &\leq C_{\mathcal{E}}(\varphi(t) + \underbrace{\mathcal{E}(\hat{t}_k, v(\hat{t}_k))}_A + 1); \end{aligned}$$

hence, using Gronwall's Lemma we obtain for  $t \in I_k$

$$\varphi(t) \leq A [e^{C_{\mathcal{E}}(t_k - t_{k-1})} - 1] \leq A [e^{C_{\mathcal{E}}\delta} - 1] \leq 2\delta A C_{\mathcal{E}},$$

which, when combined with the previous estimate, allows us to obtain

$$\int_{I_k} e^{-\lambda t} \mathcal{E}(t, v_{\delta}(t)) dt \leq \int_{I_k} e^{-\lambda t} \mathcal{E}(t, v(t)) dt + 2\delta A C_{\mathcal{E}} \int_{I_k} e^{-\lambda t} dt.$$

If we sum over  $k = 1, \dots, N$  we obtain the desired estimate (3.9), which, with the previous estimates, gives (3.1.3), concluding the proof of the Theorem.  $\square$

We are now finally ready to prove Theorem 2.1. The proof will be very similar to the one of Theorem 3.2.

*Proof of Theorem 2.1.* Let  $u_d$  be the unique minimizer of  $\mathcal{F}$  over  $\text{BV}([0, T]; Z) \cap L^2([0, T]; X_d)$  (or equivalently the unique minimizer of  $\mathcal{F}_d$  over  $\text{BV}([0, T]; Z) \cap L^2([0, T]; X)$ ). By comparison of  $u_d$  with the constant zero curve and using the uniform convexity with respect to  $\|\cdot\|_X$  of  $\mathcal{E}(t, \cdot)$ , we can obtain that there exists a constant  $C > 0$ , independent of  $d$ , such that

$$\int_0^T \|u_d(t)\|_X^2 dt \leq C.$$

But then  $(u_d)_d$  is equi-bounded in  $L^2([0, T]; X)$ , which implies precompactness with respect to the weak  $L^2_X$  topology because of the reflexivity assumption on

$(X, \|\cdot\|_X)$ . Hence, up to taking a subsequence, we can assume that  $u_d \rightharpoonup u$  in  $L^2([0, T]; X)$  for some  $u$ .

But now, using Proposition 3.2.1, we obtain that  $u$  is a minimizer of  $\mathcal{F}$  over  $BV([0, T]; Z) \cap L^2([0, T]; X_d)$ . This concludes the existence part of the result. The uniqueness of the minimizer follows from the uniform convexity of  $\mathcal{E}(t, \cdot)$ . Finally, in order to obtain the bound

$$\|\dot{u}(t)\|_X \leq \frac{C_0}{c_0}, \quad \text{for a.e. } t \in [0, T],$$

one can just observe that, again because of Theorem 3.2, the minimizers  $u_d$  satisfy

$$\|u_d(t) - u_d(s)\|_X \leq \frac{C_0}{c_0}|t - s|, \quad \text{for all } t, s \in [0, T], \quad (3.10)$$

which implies

$$\|u(t) - u(s)\|_X \leq \frac{C_0}{c_0}|t - s|, \quad \text{for all } t, s \in [0, T].$$

This follows since, at almost every point  $\tau \in [0, T]$ , for some small parameter  $\varepsilon \rightarrow 0^+$ , we have

$$(1 + o_\varepsilon(1))u_d(\tau) = \int_{\tau-\varepsilon}^{\tau+\varepsilon} u_d dr \xrightarrow{d} \int_{\tau-\varepsilon}^{\tau+\varepsilon} u dr = (1 + o_\varepsilon(1))u(\tau).$$

This concludes the proof of Theorem 2.1.  $\square$

## 3.2 Proof of Theorem (2.2)

In this Section we present our proof of Theorem 2.2. The proof is almost equal to the one of Theorem 2.1, presented in the previous Section. One can easily prove analogues of Propositions 3.1.1 and 3.2.1 for functionals  $\mathcal{F}$  which do not contain a penalization  $F$  for the initial value, but impose a Dirichlet initial condition  $u(0) = u_0$ . Indeed, the  $\Gamma$ -convergence result contained in Proposition 3.1.1 is with respect to the uniform convergence of curves, which allows the Dirichlet initial condition to pass to the limit. On the other hand, the  $\Gamma$ -convergence result of Proposition 3.2.1 is with respect to the weak  $L^2_X$  convergence; in this case the Dirichlet initial condition passes to the limit since, for all  $\varepsilon > 0$ ,

$$(1 + o_\varepsilon(1))u_d(0) = \int_0^\varepsilon u_d dr \xrightarrow{d} \int_0^\varepsilon u dr = (1 + o_\varepsilon(1))u(0).$$

The only place where a real new argument is required is the proof of the analogue of Theorem 3.1.

*Proof of analogue of Theorem 3.1.* Consider, for fixed  $\delta, \varepsilon > 0$ , the approximation functionals

$$\mathcal{F}_{\delta,\varepsilon}(u) = \int_0^T e^{-\lambda t} [L_{\delta,\varepsilon}(\dot{u}(t)) + \lambda \mathcal{E}(t, u(t))] dt,$$

where as before  $L_{\delta,\varepsilon}(v) = \sqrt{\delta^2 + \|v\|_Z^2} + \frac{\varepsilon}{2} \|v\|_X^2$ . Call  $u$  its unique minimizer among curves  $v \in H^1$ , continuous at time  $t = 0$  and such that  $v(0) = u_0$  (existence and uniqueness follows once again by the direct method and the strict convexity of  $\mathcal{E}(t, \cdot)$ ).

We would like to obtain a uniform bound (in  $\delta, \varepsilon$ ) on  $\|\dot{u}(t)\|_Z$  for all  $t \in [0, T]$ , which would allow us to conclude exactly as in the proof of Theorem 3.1.

In this case the optimality conditions for  $u$  become

$$\begin{cases} DL_{\delta,\varepsilon}(\dot{u}(t)) + D\mathcal{E}(t, u(t)) - \frac{1}{\lambda} D^2 L_{\delta,\varepsilon}(\dot{u}(t)) \ddot{u}(t) = 0, & \forall t \in [0, T], \\ u(0) = u_0, \\ DL_{\delta,\varepsilon}(\dot{u}(T)) = 0; \end{cases}$$

the major difference with the system of Theorem 3.1 is that the transversality condition at  $t = 0$  is not available anymore, and is replaced by the Dirichlet initial condition. We can still introduce the auxiliary function

$$h(v) = \langle DL_{\delta,\varepsilon}(v), v \rangle - L_{\delta,\varepsilon}(v),$$

and look for a maximum point  $t_0$  of  $h(\dot{u}(t))$  in  $[0, T]$  to obtain a bound on  $\|\dot{u}\|_X$ . If  $t_0 \in (0, T]$ , the proof contained in Theorem 3.1 still works, however, if  $t_0 = 0$  we have to find a new way to bound  $h(\dot{u}(0))$ , since we do not dispose of a transversality condition. However, by recalling

$$Dh(v) = D^2 L_{\delta,\varepsilon}(v) v,$$

we get

$$\langle D^2 L_{\delta,\varepsilon}(\dot{u}(0)) \ddot{u}(0), \dot{u}(0) \rangle \leq 0;$$

thus, the Euler-Lagrange equation and the Dirichlet condition at  $t = 0$  give us

$$\langle D\mathcal{E}(0, u_0) + DL_{\delta,\varepsilon}(\dot{u}(0)), \dot{u}(0) \rangle = \frac{1}{\lambda} \langle D^2 L_{\delta,\varepsilon}(\dot{u}(0)) \ddot{u}(0), \dot{u}(0) \rangle \leq 0,$$

from which we obtain

$$h(\dot{u}(0)) + L_{\delta,\varepsilon}(\dot{u}(0)) \leq -\langle \dot{u}(0), D\mathcal{E}(0, u_0) \rangle \leq \|\dot{u}(0)\|_Z;$$

where we have used the additional stability assumption  $\|D\mathcal{E}(0, u_0)\|_{Z^*} \leq 1$ ; but now we can use  $L_{\delta,\varepsilon}(v) \geq \|v\|_Z$  in the previous expression to get

$$\max_{t \in [0, T]} h(\dot{u}(t)) = h(\dot{u}(0)) \leq 0.$$

So, from

$$h(\dot{u}(t)) = -\frac{\delta^2}{\sqrt{\delta^2 + \|\dot{u}(t)\|_Z^2}} + \frac{\varepsilon}{2}\|\dot{u}(t)\|_X^2,$$

we obtain that, for all  $t \in [0, T]$ ,

$$-\delta + \frac{\varepsilon}{2}\|\dot{u}(t)\|_X^2 \leq h(\dot{u}(t)) \leq h(\dot{u}(0)) \leq 0 \implies \|\dot{u}(t)\|_X \leq \sqrt{\frac{2\delta}{\varepsilon}}.$$

This can be used in place of the bound (3.4) in Theorem 3.1. Thus, the remainder of the proof proceeds by simply replacing  $C_{\delta,\varepsilon,\alpha}$  by  $\frac{C_0}{c_0}$ .  $\square$



# References

- [1] Luigi Ambrosio et al. *Gradient Flows: in Metric Spaces and in the Space of Probability Measures*. Basel: Birkhäuser Basel, 2008.
- [2] Martino Bardi and Pierre Cardaliaguet. “Convergence of some Mean Field Games systems to aggregation and flocking models”. In: *Nonlinear Analysis* (2021).
- [3] Andrea Braides. “A handbook of  $\Gamma$ -convergence”. In: ed. by M. Chipot and P. Quittner. Vol. 3. North-Holland, 2006, pp. 101–213.
- [4] Pierluigi Colli and A. Visintin. “On A Class Of Doubly Nonlinear Evolution Equations”. In: *Communications in Partial Differential Equations* 15 (1990).
- [5] Ennio De Giorgi. “New problems on minimizing movements”. In: *Ennio de Giorgi: selected papers* (1993), pp. 699–713.
- [6] Annette Dumas and Filippo Santambrogio. “Optimal trajectories in  $L^1$  and under  $L^1$  penalizations”. In: *Comptes Rendus Mathématiques* (2023).
- [7] Andreas Mainik and Alexander Mielke. “Existence results for energetic models for rate-independent systems”. In: *Calculus of Variations and Partial differential equations* 22 (2005), pp. 73–100.
- [8] Alexander Mielke. “Evolution of rate-independent systems”. In: *Handbook of differential equations: Evolutionary equations*. Vol. 2. Elsevier, 2005, pp. 461–559.
- [9] Alexander Mielke and Theil Florian. “A mathematical model for rate-independent phase transformations with hysteresis”. In: *Proceedings of the Workshop on ”Models of Continuum mechanics in Analysis and Engineering”* (1999).
- [10] Alexander Mielke and Theil Florian. “On rate-independent hysteresis models”. In: *Nonlinear Differential Equations and Applications NoDEA* (2004).
- [11] Alexander Mielke and Michael Ortiz. “A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems”. In: *ESAIM: Control, Optimisation and Calculus of Variations* (2008).

- [12] Alexander Mielke and Riccarda Rossi. “Existence and uniqueness results for a class of rate-independent hysteresis problems”. In: *Mathematical Models and Methods in Applied Sciences* (2007).
- [13] Alexander Mielke and Tomas Roubicek. *Rate-Independent Systems - Theory and Application*. Springer, 2015.
- [14] Alexander Mielke and Ulisse Stefanelli. “Weighted energy-dissipation functionals for gradient flows”. In: *ESAIM: Control, Optimisation and Calculus of Variations* (2011).
- [15] Alexander Mielke, Florian Theil, and Valery Levitas. “A Variational Formulation of Rate-Independent Phase Transformations Using an Extremum Principle”. In: *Archive for Rational Mechanics and Analysis* (2002).
- [16] Riccarda Rossi et al. “Weighted Energy-Dissipation principle for gradient flows in metric spaces”. In: *Journal de Mathématiques Pures et Appliquées* (2019).
- [17] Adrien Saumard and Jon A. Wellner. “Log-concavity and strong log-concavity: A review”. In: *Statistics Surveys* 8 (2014), pp. 45–114.