Gradient flow for the Boltzmann entropy and Cheeger’s energy on time-dependent metric measure spaces

Eva Kopfer*

Abstract
We introduce notions of dynamic gradient flows on time-dependent metric spaces as well as on time-dependent Hilbert spaces. We prove existence of solutions for a class of time-dependent energy functionals in both settings. In particular in the case when each underlying space satisfies a lower Ricci curvature bound in the sense of Lott, Sturm and Villani, we provide time-discrete approximations of the time-dependent heat flows introduced in [15].

Contents
1 Introduction 2
2 Gradient flows in metric spaces 6
3 Dynamic gradient flows in time-dependent metric spaces 7
  3.1 Dynamic EDI- and EDE-gradient flows . . . . . . . . . . . . . . 11
  3.2 Dynamic EVI(K,∞)-gradient flows . . . . . . . . . . . . . . . 14
  3.3 Existence of dynamic EDI-gradient flows . . . . . . . . . . . . . 18
4 Dynamic gradient flow of the entropy 26
  4.1 Time-dependent Kantorovich metrics . . . . . . . . . . . . . . . 26
  4.2 Time-dependent Boltzmann entropy . . . . . . . . . . . . . . . 27
  4.3 Existence and Uniqueness of EDE-gradient flow for the entropy . 31
5 Dynamic gradient flows in Hilbert spaces 34
  5.1 Existence and Uniqueness . . . . . . . . . . . . . . . . . . . . . 36
6 The heat equation on time-dependent metric measure spaces 44
  6.1 Identification of the forward adjoint heat flow with the dynamic
    EDI-gradient flow for the entropy . . . . . . . . . . . . . . . . 46
  6.2 Identification of the heat flow with the dynamic gradient flow for
    Cheeger’s energy . . . . . . . . . . . . . . . . . . . . . . . . . . 51

*Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany (eva.kopfer@iam.uni-bonn.de)
1 Introduction

The heat flow in $\mathbb{R}^n$ can be interpreted as the gradient flow of the Dirichlet energy $\text{Dir}(u) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla u|^2 dx$ in the Hilbert space $L^2(\mathbb{R}^n)$. Since the seminal work of Jordan, Kinderlehrer and Otto in \cite{14} it is known that the heat flow equivalently arises as the gradient flow of the entropy functional $\text{Ent}(\rho L^n) = \int_{\mathbb{R}^n} \rho \log \rho dx$ in the space of Borel probability measures with finite second moments endowed with the $L^2$-Kantorovich distance $W_2$. This has been accomplished by using a discrete variational scheme. By now similar interpretations of the heat flow have been established in Riemannian manifolds \cite{8}, in Finsler spaces \cite{19}, in Alexandrov spaces \cite{12} and in CD($K, \infty$)-spaces \cite{11}.

The class of CD($K, \infty$)-spaces has been independently introduced by Lott and Villani in \cite{17} and Sturm in \cite{23} and consists of metric measure spaces such that the relative entropy is $K$-convex along geodesics in the $L^2$-Kantorovich space. For a Riemannian manifold this is equivalent to saying that its Ricci curvature is bounded from below by $K$. Lower Ricci curvature bounds are intimately linked to the heat equation in the sense that they can be characterized by contraction estimates of the heat flow, see e.g. \cite{27, 9}.

A similar result can be shown for time-dependent manifolds which evolve under a super-Ricci flow, see e.g. \cite{18, 13}. In \cite{24} Sturm introduced a notion for super-Ricci flows on metric measure spaces in terms of ‘dynamic convexity’ of the relative entropy. The relative entropy is now a time-dependent functional from the time-dependent $L^2$-Kantorovich space.

In this paper we present notions of gradient flows for time-dependent functionals from time-dependent metric spaces as well as time-dependent Hilbert spaces. We will prove existence for functionals which are uniformly $K$-convex using an adapted discrete variational scheme. The relative entropy and Cheeger’s energy will serve as a role model for this.

In the static CD($K, \infty$) setting the heat flow can be unambiguously defined as the gradient flow of the relative entropy or the gradient flow of the Cheeger’s functional. In the time-dependent setting the picture is less complete. However in some sense we identify the gradient flow of the relative entropy with the gradient flow of Cheeger’s energy via the heat flows in \cite{15} provided that each underlying space is RCD($K, \infty$). The condition RCD($K, \infty$) stands for CD($K, \infty$) combined with infinitesimally Hilbertian, where the latter means that Cheeger’s energy is a bilinear form. The bilinearity of Cheeger’s energy together with some regularity assumptions allow the authors in \cite{15} to prove existence and uniqueness of the heat flow in time-dependent metric measure spaces via the general theory of coercive operators. Here we identify the forward adjoint heat flow from \cite{15} with the gradient flow of the entropy provided that each space is RCD($K, \infty$), and the heat flow from \cite{15} with the gradient flow of Cheeger’s energy. In particular we provide an explicit construction of the trajectory of each heat flow.

In the following we will briefly present our main results for our model functionals.

Throughout this paper let $X$ be a topological space equipped with

- a one-parameter family $d_t$ of complete geodesic separable metrics,

- and a one-parameter family of Borel measures such that $m_t = e^{-f_t} m$ for some probability measure $m$ and suitable functions $f_t$.  

Further we assume the following
\[ |\log(d_t(x,y)/d_s(x,y))| \leq L|t-s|, \quad \text{and} \quad |f_t(x) - f_s(x)| \leq L_2^*|t-s|. \quad (1) \]

**Dynamic gradient flow of the entropy**

Given two probability measures \( \mu, \nu \) with finite second moments (with respect to any metric \( d_t \) and denoted by \( \mathcal{P}_2(X) \)) we define for every \( t \in [0,T] \) the \( L^2 \)-Kantorovich distance by
\[ W_t(\mu, \nu) = \inf \left\{ \int_{X \times X} d_t^2(x,y) \, d\pi(x,y) \ \bigg| \ \pi \text{ is a coupling of } \mu \text{ and } \nu \right\}^{1/2}. \]

The relative entropy \( S_t \) on \( \mathcal{P}_2(X) \) is defined by
\[ S_t(\mu) := \int_X \rho \log \rho \, dm_t, \]
provided that \( \mu \) has a density \( \rho \) with respect to \( m_t \). We assume that each static space \((X,d_t,m_t)\) has Ricci curvature bounded below by some \( K \in \mathbb{R} \), i.e. for each \( t \) and each \( \mu, \nu \) there exists a \( W_t \)-geodesic \((\rho_a)_{a \in [0,1]}\) connecting \( \mu \) and \( \nu \) such that
\[ S_t(\rho_a) \leq (1-a)S_t(\mu) + aS_t(\nu) - \frac{Ka(1-a)}{2}W_t^2(\mu, \nu). \quad (2) \]

This "uniform" geodesic convexity holds particularly true when the underlying spaces evolve as a super-Ricci flow introduced in [24] provided that (1) holds.

We are interested in defining a notion of gradient flows for the time-dependent entropy functional on the space of probability measures over \( X \). For this we adapt the discrete variational scheme to our dynamic setting, which we describe in the following. Fix a time step \( h > 0 \) and an initial probability measure \( \mu_0 \). Recursively define for every \( n \in \mathbb{N} \) such that \( nh \leq T \) the minimizer \( \mu_n^h \) by
\[ \mu_0^h := \mu_0, \quad \mu_n^h := \arg \min_{\nu} \left( S_{nh}(\nu) + \frac{1}{2h}W_{nh}^2(\mu_n^{h-1}, \nu) \right). \quad (3) \]

We then define a discrete trajectory as the piecewise constant interpolant \((\tilde{\mu}_t^h)_{t \in [0,T]}\) by
\[ \tilde{\mu}_0^h := \mu_0, \quad \tilde{\mu}_t^h := \mu_n^h \quad \text{if } t \in ((n-1)h, nh]. \quad (4) \]

By the direct method of the calculus of variations one can prove that there exists a unique solution to (3). Having established a sequence of minimizers \((\mu_n^h)\) the next step is to prove the existence of a curve \((\mu_t)_{t \in [0,T]}\) such that \( \tilde{\mu}_t^h \to \mu_t \) weakly as \( h \to 0 \). From the convergence of the interpolants \( \tilde{\mu}_t^h \) to some limit curve \( \mu_t \) we can deduce that \( \mu_t \) satisfies a dynamic energy dissipation inequality, in short dynamic EDI (cf. Theorem 4.7). Moreover this curve is unique (cf. Theorem 4.9).

**Theorem A.** Suppose that each \((X,d_t,m_t)\) satisfies a \( CD(K,\infty) \) condition. Then there exists an absolutely continuous curve \((\mu_t)_{t \in [0,T]} \subset \mathcal{P}_2(X)\) and a subsequence \( h_n \to 0 \) as \( n \to \infty \) such that
\[ \tilde{\mu}_t^{h_n} \to \mu_t, \text{ for every } t \in [0,T] \text{ as } n \to \infty, \]
where the convergence is to be understood in duality with bounded continuous function on $X$. Moreover this curve satisfies

$$S_t(\mu_t) + \frac{1}{2} \int_0^t |\mu_r|^2 dr + \frac{1}{2} \int_0^t |\nabla \mu_r S_r|^2(\mu_r)dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r)dr. \quad (5)$$

**Dynamic gradient flow of Cheeger’s energy**

For each $t \in [0, T]$ let us denote by $\text{Ch}_t : L^2(X, m) \to [0, \infty]$ Cheeger’s energy

$$\text{Ch}_t(u) = \frac{1}{2} \inf \left\{ \liminf_{n \to \infty} \int_X (\text{lip}_t u_n)^2 dm_t \, u_n \in \text{Lip}(X), \int_X |u_n - u|^2 dm_t \to 0 \right\},$$

where $\text{lip}_t u$ denotes the local slope. By making use of the minimal relaxed gradient $|\nabla_t u|_*$ (\cite[Definition 4.2]{[3]}), this functional admits an integral representation

$$\text{Ch}_t(u) = \frac{1}{2} \int_X |\nabla_t u|^2_* dm_t,$$

set equal to $+\infty$ if $u$ has no relaxed slope. The subdifferential $D_{-t} \text{Ch}_t(u)$ of $\text{Ch}_t$ at some $u \in \text{Dom}(\text{Ch}_t)$ is the set of $v \in L^2(X, m)$ such that

$$\text{Ch}_t(w) - \text{Ch}_t(u) \geq \langle v, w - u \rangle_t \quad \forall w \in L^2(X, m),$$

where we set $\langle v, w \rangle_t = \int vw dm_t$. We prove the existence of a dynamic gradient flow for $(\text{Ch}_t)$, cf. Theorem 6.10, via a discrete variational scheme similar to (3). Since each $\text{Ch}_t$ already defines a convex functional, we do not need any curvature condition on the space.

**Theorem C.** Let $\bar{u} \in \text{Dom}(\text{Ch})$. Then there exists a unique gradient flow for $(\text{Ch}_t)_{t \in [0, T]}$ starting in $\bar{u}$, i.e. an absolutely continuous curve $u : [0, T] \to \text{Dom}(\text{Ch})$ solving

$$\partial_t u_t \in -D_{-t} \text{Ch}_t(u_t) \quad \text{for a.e. } t \in (0, T) \quad (6)$$

and $\lim_{t \to 0} u_t = \bar{u}$.

Let us emphasize here that many facts from the static theory of gradient flows hold no longer true in the time-dependent framework. For example it is no longer valid that we have convergence of minimizers (Example 2), or ‘EVI’ implies ‘EDE’ (Example 5), or a ‘minimal selection principle’ (Example 6).

**Heat flows on time-dependent metric measure spaces**

In the static $\text{CD}(K, \infty)$ setting it is a well-known fact that the heat flow can be unambiguously defined as the gradient flow of the entropy or as the gradient flow of Cheeger’s energy. Here we provide a similar result using the heat and the adjoint heat flow introduced in [15].

Let us assume that each $(X, d_t, m_t)$ is infinitesimally Hilbertian so that $\mathcal{E}_t := 2\text{Ch}_t$ becomes a Dirichlet form with generator $\Delta_t$. The heat flow solving $\partial_t u = \Delta_t u$ as well as the adjoint heat flow solving $\partial_s v_s = -\Delta_t v_s + f_s v_s$ has to be understood in a weak distributional sense. If we assume further that $|f_t(x) - f_t(y)| \leq C d_t(x, y)$ the existence and uniqueness of both flows are ensured by the
general theory of time-dependent coercive operators on some fixed Hilbert space $L^2(X,m)$. The flows denoted by $P_{t,s}u$ and $P_{t,s}^*v$ are adjoint in the sense that

\[ \int P_{t,s}u \cdot v \, dm = \int u P_{t,s}^*v \, dm. \]

Let us remark that many properties which are apparent for the static heat flow on metric measure spaces hold no longer true for the time-dependent version, or require some extra effort, e.g. semigroup and generator commute or the semigroup maps $L^2$ into the domain of the generator.

The following theorem states that the trajectory of the adjoint heat flow parametrized forwards in time coincides with the trajectory of the gradient flow of the relative entropy. For the precise statement see Theorem 6.7.

**Theorem B.** Suppose that each $(X,d_t,m_t)$ satisfies an RCD($K,\infty$) condition. Let $(\mu_t)_{t \in [0,T]}$ be a continuous curve in $\mathcal{P}_2(X)$. Then the following are equivalent:

1. $(\mu_t)$ is a gradient flow for the relative entropy,
2. $(\mu_t)$ is given by $\mu_t(dx) = \rho_t(x)m_t(dx)$, where $(\rho_t)$ is a solution to the adjoint heat equation
\[
\partial_t \rho_t(x) = \Delta_t \rho_t(x) + \rho_t(x) \partial_t f_t(x).
\]

In a second step we identify the gradient flow with the heat flow, see Theorem 6.11. Hence under the RCD($K,\infty$) assumption the gradient flow of the relative entropy and the gradient flow of Cheeger’s energy are connected by the heat and the adjoint heat flow.

**Theorem D.** Suppose that each $(X,d_t,m_t)$ is infinitesimally Hilbertian. Let $(u_t)$ be a continuous curve in $L^2(X,m)$. Then the following are equivalent:

1. $(u_t)$ is a gradient flow for Cheeger’s energy,
2. $(u_t)$ is a solution to the heat equation
\[
\partial_t u_t = \Delta_t u_t.
\]

**Related work**

Gradient flow formulations for time-dependent functionals similar to (5) and (6) have been considered recently. In [21], Rossi, Mielke and Savaré analyze doubly nonlinear evolution equations on a reflexive Banach space $V$

\[
D^- \psi(\partial_t u_t) + D^- E_t(u_t) \geq 0 \quad \text{in } V^* \quad \text{for a.e. } t \in (0,T).
\]

They propose a formulation of a similar form as (5) and prove existence of solutions by making use of a discrete variational scheme. In [10], Ferreira and Valencia-Guevara consider the Fokker-Planck equation

\[
\partial_t \rho = \kappa(t) \Delta \rho + \nabla \cdot (\nabla V(t,x)\rho) \quad \text{on } \mathbb{R}^d \times [0,\infty),
\]

for some fixed non-increasing absolutely continuous function $\kappa: [0,\infty) \to [0,\infty)$, and potential $V: [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$. They show that the solution can be obtained as the gradient flow of some specified time-dependent energy functional. They also use a version of the discrete variational scheme similar to [21] to obtain existence. In both papers the authors do not treat time-dependent metric measure spaces as done here.
Organization of the article

In Section 2 we briefly recall the concept of gradient flows in metric spaces. In Section 3 we introduce the notion of dynamic EDI- and dynamic EDE-gradient flows on time-dependent metric spaces \((X,d_t)_{t \in [0,T]}\) satisfying (1). Moreover we introduce the notion of dynamic EVI-gradient flows and show that it implies EDE restricted to a suitable class of energy functionals. We show existence of dynamic EDI-gradient flows for a class of energy functionals \(E: [0,T] \times X \to (-\infty, +\infty]\) and give sufficient conditions for the existence of EDE-gradient flows. In Section 4 we apply the results from Section 3 and prove existence and uniqueness of dynamic EDI-gradient flows in time-dependent metric measure spaces \((X,d_t,m_t)_{t \in [0,T]}\) for the time-dependent entropy functional \(S: [0,T] \times \mathcal{P}_2(X) \to (-\infty, +\infty]\). In Section 5 we consider dynamic gradient flows in the form of (6) on time-dependent Hilbert spaces \((H,\langle \cdot, \cdot \rangle_t)_{t \in [0,T]}\) and show that they imply EVI. We prove existence and uniqueness of such gradient flows for a class of energy functionals \(E: [0,T] \times H \to [0, +\infty]\). In Section 6 we recall the concept of heat equation on time-dependent metric measure spaces introduced in [15]. We identify the dynamic EDI-gradient flow of the entropy with the forward adjoint heat flow. We apply the results from Section 5 and directly obtain existence and uniqueness of a dynamic gradient flow for Cheeger’s energy and identify it with the heat flow.

Acknowledgement
I would like to thank my supervisor Karl-Theodor Sturm for supporting me. I also thank Peter Gladbach for helpful comments and suggestions on this paper.

2 Gradient flows in metric spaces

We briefly recall the notions of gradient flows on metric spaces \((X,d)\). A curve \(x: [a,b] \to X\) is said to belong to \(AC^p([a,b];X)\) for \(1 \leq p \leq \infty\), if there exists \(g \in L^p([a,b])\) such that
\[
d(x_s,x_t) \leq \int_s^t g(r)dr\quad \text{for every } a \leq s \leq t \leq b. \tag{7}
\]

The metric speed of \(x\), defined by
\[
|\dot{x}_t| := \lim_{h \to 0} \frac{d(x_{t+h},x_t)}{h},
\]
exists for a.e. \(t \in (a,b)\), is of class \(L^p(a,b)\) and is the smallest function such that (7) holds, see e.g. [11 Theorem 1.1.2].

Given \(E: X \to (-\infty, +\infty]\) we define the slope \(|\nabla E|(x)\) at \(x\) by
\[
|\nabla E|(x) := \limsup_{y \to x} \frac{(E(x) - E(y))^+}{d(x,y)}.
\]

We now are ready to give three possible definitions of gradient flows in a metric framework, cf. [2].
Definition 2.1. 1. An absolutely continuous curve \((x_t) \subset X\) is a EDI-gradient flow if it satisfies the following Energy Dissipation Inequality

\[
E(x_s) + \frac{1}{2} \int_s^t |\dot{x}_r|^2 \, dr + \frac{1}{2} \int_s^t |\nabla E|^2(x_r) \, dr \leq E(x_t) \quad \forall s < t. \tag{8}
\]

2. An absolutely continuous curve \((x_t) \subset X\) is an EDE-gradient flow if it satisfies the following Energy Dissipation Equality

\[
E(x_s) + \frac{1}{2} \int_s^t |\dot{x}_r|^2 \, dr + \frac{1}{2} \int_s^t |\nabla E|^2(x_r) \, dr = E(x_t) \quad \forall s < t. \tag{9}
\]

3. An absolutely continuous curve \((x_t) \subset X\) is an EVI-gradient flow (with respect to \(\lambda \in \mathbb{R}\)) if it satisfies the following Evolution Variation Inequality

\[
E(x_t) + \frac{1}{2} \partial_t d^2(x_t, y) + \frac{\lambda}{2} d^2(x_t, y) \leq E(y) \quad \text{for a.e. } t \in [0, T], \forall y \in X. \tag{10}
\]

It holds in general that EVI implies EDE, and EDE implies trivially EDI. If the underlying space is a Hilbert space and the energy functional is convex, all three formulations are equivalent. Moreover we can characterize the flow in terms of the subdifferential by

\[
\dot{x}_t \in -D^{-}E(x_t), \tag{11}
\]

where \(D^{-}E(x)\) consists of all \(v \in X\) such that

\[
E(x) + \langle v, y - x \rangle \leq E(y) \quad \forall y \in X.
\]

In this paper we are interested in finding substitutions for formulations of the form (8) and (9), where the metric as well as the functional varies in time. A formulation in the sense of (10) has already been introduced in \([15]\). Moreover, in the Hilbert space case, we study the time-dependent counterpart of relations of the form (11).

3 Dynamic gradient flows in time-dependent metric spaces

In the sequel we fix a one-parameter family of complete geodesic metric spaces \((X, d_t)\) indexed by \(t \in [0, T]\). We always assume that the map \(t \to \log d_t(x, y)\) is Lipschitz continuous, i.e. there exists a constant \(L\) such that

\[
|\log(d_t(x, y)/d_s(x, y))| \leq L|t - s|. \tag{12}
\]

We give a simple example for this setting.

Example 1. Let \(M\) be a smooth closed manifold equipped with a smooth family of Riemannian metrics \((g_t)\) evolving under a Ricci flow, i.e.

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]

\[
\frac{1}{2} \partial_t g_t = -\mathrm{Ric}(g_t),
\]
Lemma 3.2. For any curve \( \gamma \) where \( |\gamma| \leq L \), we have metric equivalence
\[
|\partial_t \log g_t(v,v)| \leq L.
\]
This implies that (12) holds for the geodesic distances \( (d_t) \).

The metric speed

**Definition 3.1.** Let \([0,T]\) \( \ni t \mapsto x_t \in X \) be a curve. We say that \((x_t) \in AC^p([0,T]; X)\), for \( p \in [1,\infty] \), if for any (and thus for all) \( t^* \in [0,T] \) there exists a function \( g \in L^p(0,T) \) such that
\[
d_{t^*}(x_t, x_s) \leq \int_t^s g(r)dr \quad \forall 0 \leq t \leq s \leq T.
\]

We define the length of a curve \( x : [0,T] \to X \) to be
\[
L_x(t) = \lim_{h \to 0} \sup \left\{ \sum_{i=1}^n d_i(x_{t_i}, x_{t_{i+1}}) : 0 = t_1 < \ldots < t_n = t, t_{j+1} - t_j \leq h \right\}.
\]

It is a direct consequence of the definition of \( L_x(t) \) that if \( x_n \to x \) pointwise as \( n \to \infty \) we have \( L_x(t) \leq \lim \inf_{n \to \infty} L_{x_n}(t) \) for every \( t \in [0,T] \).

Note that \( L_x \) is absolutely continuous as soon as \( x \) is and hence we may define the instantaneous speed of the curve as the derivative of its length.
\[
|\dot{x}|_t := \dot{L}_x(t).
\]

**Lemma 3.2.** For any curve \( x \in AC^p([0,T]; X) \) the function \( t \mapsto |\dot{x}|_t \) is in \( L^p(0,T) \), and for almost every \( t \in (0,T) \)
\[
|\dot{x}|_t = \lim_{s \to t} \frac{d_t(x_s, x_t)}{|s-t|} = |\dot{x}|_t.
\]

**Proof.** If we show the second assertion the first assertion is an easy consequence of (12). Let \( x \) be an absolutely continuous curve and choose an arbitrary partition \( s = t_1 < t_2 < \ldots < t_{N+1} = t \). Then we find
\[
d_t(x_t, x_s) \leq \sum_{i=1}^{N} d_t(x_{t_i}, x_{t_{i+1}}) + C \sum_{i=1}^{N} |t-t_i|d_t(x_{t_i}, x_{t_{i+1}})
\leq \sum_{i=1}^{N} d_t(x_{t_i}, x_{t_{i+1}}) + C|t-s| \int_s^t g(r)dr,
\]
where we used (12) and \( g \) is some \( L^p \) function. Hence we may estimate
\[
d_t(x_t, x_s) \leq L(t) - L(s) + C|t-s| \int_s^t g(r)dr.
\]
Dividing by \( |t-s| \) and letting \( s \to t \) we deduce
\[
\limsup_{s \to t} \frac{d_t(x_t, x_s)}{|t-s|} \leq \dot{L}_x(t) \text{ for almost every } t.
\]
We show the other inequality by contradiction. Fix $\eta > 0$ and consider the set of points

$$F = \left\{ t : \liminf_{s \to t} \left( \frac{d_t(x_s, x_t)}{|s - t|} - \frac{1}{|s - t|} \int_s^t \dot{L}_x(r) dr \right) < -\eta \right\}.$$  

We assume that the Lebesgue outer measure $L^*(F) > 0$. Fix $\delta > 0$ and cover the set $F$ with intervals

$$F := \bigcup_{t \in F} (t - \delta_t, t + \delta_t),$$  

such that

$$d_t(x_s, x_t) < \int_s^t \dot{L}_x(r) dr - |t - s|\eta/2 \quad (13)$$  

for all $t \in F$ and some $s \in (t - \delta_t, t + \delta_t)$. From the Besicovitch covering theorem [6, Theorem 5.8.1] it follows that there exists a constant $N$ and a subcollection $\mathcal{F}_1, \ldots, \mathcal{F}_N$ each consisting of at most countably many disjoint intervals $B$ such that

$$F \subset \bigcup_{i=1}^N \bigcup_{B \in \mathcal{F}_i} B.$$  

Since the outer measure of $F$ is strictly positive we can find a family $\mathcal{F}_j$ of at most countably many disjoint intervals denoted by $\mathcal{F}_j = \{(t_i - \delta_i, t_i + \delta_i), i \in I\}$ such that $\mathcal{L}^1(\bigcup_{B \in \mathcal{F}_j} B) \geq \frac{1}{N} L^*(F) > 0$.

We define a curve $x^\delta : [0, T] \to X$ in the following way

$$x^\delta_t = \begin{cases} 
  x_{t_i} & \text{if } t \in (t_i, t_i + \delta_i) \\
  x_t & \text{else.}
\end{cases}$$

Note that this curve is not continuous but still its length is finite. Further we observe that $x^\delta_t$ converges to $x_t$ pointwise as $\delta$ goes to 0 and hence

$$\liminf_{\delta \to 0} L_{x^\delta}(T) \geq L_x(T).$$

It suffices to show that

$$L_{x^\delta}(T) \leq L_x(T)(1 + L\delta) - \frac{\eta}{2} \mathcal{L}^1 \left( \bigcup_{i \in I} (t_i, t_i + \delta_i) \right), \quad (14)$$

since then

$$L_x(T) \leq \liminf_{\delta \to 0} L_{x^\delta}(T) \leq L_x(T) - \frac{\eta}{4N} \mathcal{L}^*(F) < L_x(T),$$

which is clearly a contradiction. Hence for the outer measure it must hold $\mathcal{L}^*(F) = 0$ and therefore already $\mathcal{L}^1(F) = 0$. Since $L_x$ is absolutely continuous we conclude

$$\liminf_{s \to t} \frac{d_t(x_s, x_t)}{|s - t|} \geq \dot{L}_x(t) \text{ for almost every } t \in [0, T].$$
It remains to show (13). Take a partition \((p_j)_{j=1}^m\) of \([0, T]\), with \(0 < p_{j+1} - p_j \leq h\) and \(h << \delta\). Consider the points near the boundary of \((t_i, t_i + \delta_i)\)

\[ j_i^\leq := \max\{j | p_j \leq t_i, 1 \leq j \leq m\}, \quad j_i^\geq := \min\{j | p_{j+1} \geq t_i + \delta_i, 1 \leq j \leq m\}. \]

Since \(x\) is absolutely continuous we can estimate

\[ d_{p_j^\leq}(x_{p_j^\leq, i}, x_{p_j^\leq, i+1}) = d_{p_j^\leq}(x_{p_j^\leq, i}, x_{p_j^\leq, i+1}) \leq \int_{p_j^\leq}^{p_j^\leq + h} g(r) dr, \]

where \(g \in L^p(0, T)\). Applying (13), (12) and again the absolute continuity we obtain

\[ d_{p_j^\leq}(x_{p_j^\leq, i}, x_{p_j^\leq, i+1}) \leq d_{p_j^\leq}(x_{t_i, x_{t_i + \delta_i}}) + \int_{t_i}^{t_i + \delta_i + h} g(r) dr \]

\[ \leq \int_{t_i}^{t_i + \delta_i + h} (\frac{\dot{L}_x(r) - \eta/2}{1 + \dot{L}_x(r)} + \int_{t_i}^{t_i + \delta_i + h} g(r) dr). \]

Taking the supremum over all partitions \((p_j)\) and letting \(h \to 0\) we can estimate the length of the curve \(x^s\)

\[ L_{x^s}(T) \leq \int_{(0, T) \setminus \cup_i (t_i, t_i + \delta_i)} \dot{L}_x(r) dr + \sum_i \int_{t_i}^{t_i + \delta_i} (\dot{L}_x(r) - \eta/2) dr (1 + C^* \delta_i) \]

\[ \leq \int_{(0, T)} \dot{L}_x(r) dr (1 + L \delta) - \eta / 2 L^1 \left( \bigcup_i (t_i, t_i + \delta_i) \right), \]

which proves (13).

\[ \square \]

The slope

**Definition 3.3.** Let \(E : [0, T] \times X \to (-\infty, +\infty]\) and \(s, t \in [0, T], x \in X\) with \(E_t(x) < \infty\). Then the slope \(\nabla_s E_t(x)\) of \(E_t\) with respect to \(d_s\) is given by

\[ |\nabla_s E_t(x)| = \lim_{y \to x} \sup \frac{|E_t(x) - E_t(y)|^+}{d_s(x, y)} = \lim_{y \to x} \max \left\{ \frac{E_t(x) - E_t(y)}{d_s(x, y)}, 0 \right\}. \]

We mainly deal with the case \(t = s\) in the definition of the slope. We estimate the deviation of the \(d_s\) slope from the \(d_s\) slope in the following lemma.

**Lemma 3.4.** Let \(s, t \in [0, T]\) and \(x \in X\) such that \(E_t(x) < \infty\) and \(|\nabla_s E_t(x)| < \infty\). Then \(\nabla_t E_t(x) < \infty\) and

\[ ||\nabla_t E_t(x) - |\nabla_s E_t(x)| \leq L |t - s||\nabla_s E_t(x)\|. \]

**Proof.** This follows from (12) and \(\log r \leq r - 1\) and \(\log(r^{-1}) \geq 1 - r\) respectively. \[ \square \]
3.1 Dynamic EDI- and EDE-gradient flows

Let us first motivate the definition of dynamic EDI-gradient flows by considering a Hilbert space $X$ endowed with a family of scalar products $(\langle \cdot, \cdot \rangle_t)_{t \in [0, T]}$ depending smoothly on $t$. Let $E_t : X \to \mathbb{R}$ be a $C^1$ functional also smoothly depending on time. In this setting we understand a gradient flow as a curve solving

$$\dot{x}_t = -\nabla_t E_t(x_t). \quad (15)$$

Let us observe that (15) can be rewritten as

$$\frac{d}{dt} E_t(x_t) \leq -\frac{1}{2} |\nabla_t E_t|_t^2(x_t) - \frac{1}{2} |\dot{x}_t|_t^2 + (\partial_t E_t)(x_t), \quad (16)$$

where $(\partial_t E_t)(x_t)$ stands for $\frac{d}{ds} E_s(x_t) \bigg|_{s=t}$. Indeed, along any differentiable curve it holds

$$\frac{d}{dt} E_t(x_t) = \frac{d}{ds} E_s(x_t) \bigg|_{s=t} + \langle \nabla_t E_t(x_t), \dot{x}_t \rangle_t \geq \frac{d}{ds} E_s(x_t) \bigg|_{s=t} - \frac{1}{2} |\nabla_t E_t|_t^2(x_t) - \frac{1}{2} |\dot{x}_t|_t^2,$$

and we have equality if and only if (15) holds. The functional’s dependence on the time variable leads to a “drift” of the gradient flow, i.e. in some sense the gradient flow does not follow the steepest descent. To illustrate this we give an example about the asymptotic behavior.

**Example 2.** Let $X = \mathbb{R}$ and $d_t(x, y) = |x - y|$ for $t \in [0, \infty)$ and $x, y \in \mathbb{R}$. We consider the energy $E_t(x) = (x - t)^2$ and the curve $x_t = \frac{1}{t} e^{-2t} + t - \frac{1}{t}$. Note that

$$\dot{x}_t = -e^{-2t} + 1 = -2(x_t - t) = -\partial_t E_t(x),$$

and hence $(x_t)$ is a gradient flow. A well-known fact in the theory of gradient flows is that for strictly convex functionals the gradient flow converges to the minimum of the functional as $t \to \infty$, see e.g. [1, Theorem 3.1(v)]. In our case the minima depend on time and are given by $x_t^{\min} = t$. Hence

$$|x_t - x_t^{\min}| = \frac{|e^{-2t} - 1|}{2},$$

which obviously does not converge to 0 as $t \to \infty$.

Let us now come back to our original family of complete, separable, geodesic, metric spaces $(X, d_t)$ such that (12) holds true. We call a measurable functional $E$ on $[0, T] \times X$ admissible if it satisfies the following assumptions.

**A1** The domain $\text{Dom}(E_t) := \{ x \in X | E_t(x) < \infty \}$ is time-independent and nonempty.

**A2** For each $t \in [0, T]$, $x \mapsto E_t(x)$ is uniformly bounded from below.

**A3** For each $t \in [0, T]$, $x \mapsto E_t(x)$ is lower semicontinuous.
A4 The map \( t \mapsto E_t(x) \) is uniformly Lipschitz continuous, i.e. there exists a constant \( L^* \) such that
\[
|E_t(x) - E_s(x)| \leq L^*|t - s| \quad \forall t, s \in [0, T], x \in \text{Dom}(E),
\]
and the set of differentiability points of the map \( t \mapsto E_t(x) \) can be chosen regardless of \( x \in X \) as soon as \( x \in \text{Dom}(E) \).

Note that the Lipschitz continuity of the map \( t \mapsto E_t(x) \) provides a.e. differentiability in \( t \) for every fixed \( x \). But this is not enough to get a meaningful expression in (16), since we may have that for some absolutely continuous curve \((x_t), t \mapsto E_t \) is not differentiable at \( x_t \) for every \( t \). To circumvent this problem we suppose that the set of differentiability points can be chosen independent of \( x \), cf. [10][21]. To illustrate this we give the following example, which has also been discussed in [21].

Example 3. Let \( X = \mathbb{R} \) and \( d_t(x, y) = |x - y| \) for every \( t \in [0, T] \), \( x, y \in \mathbb{R} \). Consider the following energy functional \( E : [0, T] \times \mathbb{R} \rightarrow [0, \infty) \) given by
\[
E_t(x) = |x - t|.
\]
Then the map \( t \mapsto E_t(x) \) is clearly Lipschitz continuous with well-defined derivatives \( \partial_t E_t(x) \) as long as \( t \in [0, T] \setminus \{x\} \). If we choose the curve \((x_t)_{t \in [0, T]} \in C^\infty([0, T]) \) by setting \( x_t = t \), the map \( s \mapsto E_s(x_t) \) is not differentiable at any \( t \in [0, T] \). Indeed, for every \( t \in [0, T] \) the right derivative \( \partial_x E_s(x_t)|_{s=t^+} \) equals 1, while the left derivative \( \partial_x E_s(x_t)|_{s=t^-} \) equals -1.

Definition 3.5. We call a locally absolutely continuous curve \( x : [0, T] \rightarrow X \) a dynamic EDI-gradient flow for an admissible functional \( E : [0, T] \times X \rightarrow (\mathbb{R} \cup \{\infty\}) \), if for every \( t \in [0, T] \)
\[
E_t(x_t) + \frac{1}{2} \int_0^t |\dot{x}_t|^2 dr + \frac{1}{2} \int_0^t \|
abla_x E_r\|^2(x_r) dr \leq E_0(x_0) + \int_0^t (\partial_x E_r)(x_r) dr,
\] (17)
where we used the shorthand notation \( (\partial_x E_t)(x_t) = \frac{d}{dt} E_t(x_t)|_{t=1} \). We call a locally absolute continuous curve \( x : [0, T] \rightarrow X \) a dynamic EDE-gradient flow for an admissible functional \( E : [0, T] \times X \rightarrow (\mathbb{R} \cup \{\infty\}) \), if for every \( t \in [0, T] \)
\[
E_t(x_t) + \frac{1}{2} \int_0^t |\dot{x}_t|^2 dr + \frac{1}{2} \int_0^t \|
abla_x E_r\|^2(x_r) dr = E_0(x_0) + \int_0^t (\partial_x E_r)(x_r) dr,
\] (18)
Clearly, (18) implies (17). In the following we want to give sufficient conditions for the other implication.

Definition 3.6. We say that the above mentioned functional \( E \) is \( K \)-convex for \( K \in \mathbb{R} \), if for every \( t \in [0, T] \) and for any pair of points \( x, y \in \text{Dom}(E) \) there exists a \( d_t \)-geodesic \((\gamma_a)_{a \in [0, 1]}\) connecting \( x \) and \( y \) such that for all \( a \in [0, 1] \)
\[
E_t(\gamma_a) \leq (1 - a)E_t(\gamma_0) + aE_t(\gamma_1) - K \frac{a(1-a)}{2} d_t^2(\gamma_0, \gamma_1).
\] (19)
The convexity assumption allows us to reformulate the slope
\[
|\nabla_x E_t|(x) = \sup_{y \neq x} \left[ \frac{E_t(x) - E_t(y)}{d_t(x, y)} + \frac{K}{2} d_t(x, y) \right]^+,
\] (20)
with $K^- := \max\{0, -K\}$, cf. \cite[Theorem 2.4.9]{11}.

The next proposition can be thought of as a weak chain rule in the sense of \cite[Proposition 3.19]{10}. The convexity of the functional plays an important role in the proof of this result. Unlike in the static case we additionally have to impose a condition on the difference quotients of the functionals, cf. \cite[Theorem 5.4]{10}.

**Proposition 3.7.** Let $E: [0, T] \times X \to (-\infty, +\infty]$ be a $K$-convex admissible functional. Moreover assume that for almost every $t \in [0, T]$

$$\liminf_{n \to \infty} \frac{E_{t_n}(x_n) - E_t(x_n)}{t_n - t} \geq \partial_t E_t(x), \quad \text{if } t_n \searrow t, \quad x_n \xrightarrow{d} x \text{ as } n \to \infty. \quad (21)$$

Then for every locally absolutely continuous curve $(x_t) \subset \text{Dom}(E)$, the function $t \mapsto E_t(x_t)$ is absolutely continuous and it holds

$$E_t(x_t) - E_s(x_s) \geq \int_s^t \left( \partial_v E_v \right)(x_v) \, dv - \int_s^t |x_v| \nabla_x E_v(x_v) \, dv, \quad s < t. \quad (22)$$

In particular, if $(x_t)$ is a dynamic EDI-gradient flow, it is a dynamic EDE-gradient flow as well.

**Proof.** In view of \cite[Lemma 1.1.4(a)]{12} we can find an increasing and absolutely continuous map $s: [0, T] \to [0, T']$, whose inverse $t$ is Lipschitz. The reparametrization $x_{s(t)} := (t) \text{satisfies } |\dot{x}_s| \leq 1 \text{ for almost every } s \in [0, T] \text{ with respect to some fixed metric } d_{t^*}. \text{ Notice that it is sufficient to prove that } s \mapsto E_t(s(\dot{x}_s)) = \varphi(s) \text{ is absolutely continuous, as then } E_t(x_t) = E_t(x_{s(t)}) \text{ is absolutely continuous and for almost every } t \in [0, T]$

$$\frac{d}{dt} E_t(x_t) = \lim_{h \to 0} \frac{E_{t+h}(x_{t+h}) - E_t(x_t)}{h} \geq \liminf_{h \to 0} \frac{E_{t+h}(x_{t+h}) - E_t(x_t)}{h} + \liminf_{h \to 0} \frac{E_t(x_{t+h}) - E_t(x_t)}{h} \frac{d_t(x_{t+h}, x_t)}{d_t(x_{t+h}, x_t)}$$

$$\geq \partial_t E_t(x_t) - \limsup_{h \to 0} \frac{E_t(x_{t+h}) - E_t(x_t)}{h} \frac{d_t(x_{t+h}, x_t)}{d_t(x_{t+h}, x_t)}$$

where we used (21) in the third inequality. After integration we obtain (22).

In view of the convexity of $E$ we may use the representation formula of the slope (20) and write using $(a^+)^\leq (a + b)^\leq + b^-$ and the Lipschitz property of the functional

$$\varphi(s_1) - \varphi(s_0) \leq |\nabla_x E_t(s_1)(\dot{x}_{s_1})d_{t(s_1)}(\dot{x}_{s_1}, \hat{x}_{s_0})| + K^-|d_{t(s_1)}(\dot{x}_{s_1}, \hat{x}_{s_0})| + L^*|s_1 - s_2|$$

$$\leq \left( |\nabla_x E_t(s_1)(\dot{x}_{s_1})| + \frac{K^-}{2} \right) e^{C|s_1 - s_0|} + L^*|s_1 - s_0|, \quad (23)$$

where $D$ is the finite diameter of the image $\{\hat{x}_s\}$ with respect to $d_{t^*}$. Changing the roles of $s_0$ and $s_1$ yields

$$|\varphi(s_1) - \varphi(s_0)| \leq \left( |\nabla_x E_t(s_1)(\dot{x}_{s_1}) + |\nabla_x E_t(s_0)(\dot{x}_{s_0}) + \frac{K^-}{2} \right) e^{C|t - s|} + L^*|t - s|.$$
Applying [2, Lemma 1.2.6] we conclude that the map \( s \mapsto \varphi(s) \) is in the Sobolev space \( W^{1,1}(0, T') \). To prove absolute continuity we simply check that it coincides with its continuous representative. We already know that \( s \mapsto \varphi(s) \) is lower semicontinuous and therefore continuity follows if we show
\[
\limsup_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(s + r) dr \leq \varphi(s) \quad \forall s \in (0, T').
\]
This can be seen by applying (23) once more and we get
\[
\limsup_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(s + r) - \varphi(s) dr
\]
\[
\leq \limsup_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \left( \nabla t_{s+r} E_t(s+r)|\tilde{x}_{s+r}| + \frac{K^2}{2} D^2 \right) e^{-C|r|} + L^*e^r dr
\]
\[
\leq \limsup_{\varepsilon \to 0} \frac{1}{2} \int_{-\varepsilon}^{\varepsilon} \left( \nabla t_{s+r} E_t(s+r)|\tilde{x}_{s+r}| + \frac{K^2}{2} D^2 \right) e^C + L^* dr = 0.
\]

3.2 Dynamic EVI\((K, \infty)\)-gradient flows

Let us now come to the dynamic version of EVI\((K, \infty)\)-gradient flows introduced in [15].

Definition 3.8. For \( s, t \in [0, T] \) and an absolutely continuous curve \((x_a)_{a \in [0,1]}\), we define the action
\[
\mathcal{A}_{s,t}(x) = \lim_{h \to 0} \sup \left\{ \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} d^2_{\vartheta(a_{i-1})}(x_{a_{i-1}}, x_{a_i}) \right\},
\]
where \( \vartheta : [0,1] \to [0, \infty) \) denotes the linear interpolation with \( \vartheta(0) = s \) and \( \vartheta(1) = t \). For two points \( x^0, x^1 \in X \) we define
\[
d^2_{s,t}(x^0, x^1) = \inf\{ \mathcal{A}_{s,t}(x) | x : [0,1] \to X \text{ absolutely continuous}, x_0 = x^0, x_1 = x^1 \}.
\]

Note that using the definition of the metric speed we obtain for the action the more intuitive expression, cf. [15, Proposition 7.2],
\[
\mathcal{A}_{s,t}(x) = \int_0^1 |\dot{x}_a|^2_{\vartheta(a)} da.
\]

We understand \( d_{s,t}(x, y) \) as “dynamic distance” between the points \( x \) and \( y \). In the next example we give a formula for \( d_{s,t} \) in the special case \((\mathbb{R}^n, d_t)\), where \( d^2_t(x, y) = \lambda t|x - y|^2 \).

Example 4. Consider \( \mathbb{R} \) endowed with the distance \( d^2_t(x, y) = \lambda t|x - y|^2 \) for some \( \lambda > 0 \). Then for each \( x, y \in \mathbb{R} \) and \( s < t \)
\[
d^2_{s,t}(x, y) = \frac{\lambda(t-s)}{\log t - \log s} |x - y|^2.
\]
Indeed, we have to compute the infimum of \( \int_0^1 (s + \alpha(t - s)) (\gamma_0) \) among all \( (\gamma_0)_{\alpha \in [0,1]} \) such that \( \gamma_0 = x, \gamma_1 = y \). From the Euler-Lagrange equation
\[
0 = \frac{d}{dt}((s + \alpha(t - s)) \gamma_0)
\]
we deduce that
\[
\gamma_0 = \frac{t - s}{\log t - \log s} \frac{y - x}{s + \alpha(t - s)}
\]
and hence
\[
\int_0^1 (s + \alpha(t - s)) (\gamma_0) = \frac{t - s}{\log t - \log s} |y - x|^2,
\]
which proves the claim. The formula for \( d_s \) on \( \mathbb{R}^n \) can be shown analogously by considering each component separate. Let us remark that the factor \( \theta(s, t) := (t - s)/(\log t - \log s) \) is also known as the logarithmic mean.

Besides the above example we have in general \( d_{s,t}(x, y) \neq d_{s,t}(y, x) \) as soon as \( x \neq y \). However, it clearly holds \( d_{s,t}(x, y) = d_{t,s}(y, x) \), \( d_{t,t}(x, y) = d_{t}(y, x) \) and \( d_{s,t}(x, x) = 0 \).

We will use the following notation: \( \partial^+_t u(t) := \limsup_{a \to t} \frac{u(t) - u(s)}{t - s} \).

**Definition 3.9.** Let \( E : [0, T] \times X \to (-\infty, \infty] \) be a lower semicontinuous functional in \( X \). An absolutely continuous curve \((x_t)_{0 \leq t \leq T}\) is called dynamic \( E \)-gradient flow for \( E \) if for all \( t \in (0, T) \) and all \( y \in \text{Dom}(E_t) \)
\[
\frac{1}{2} \partial^+_s d_{s,t}^2(x_s, y) \bigg|_{s=t} + \frac{K}{2} d_{t}^2(x_t, y) \leq E_t(y) - E_t(x_t).
\]

We say that the gradient flow \((x_t)_{0 \leq t \leq T}\) starts in \( x' \in X \) if \( \lim_{t \to 0} x_t = x' \).

We show uniqueness of dynamic \( E(V, \infty) \) flows by proving a contraction estimate. This estimate involves the logarithmic Lipschitz control \( L \) from [12]. For an estimate without this control see Theorem 7.7 in [15].

**Lemma 3.10.** The following holds true.

1. Suppose that \((x_t)\) is a \( E(V, \infty) \)-gradient flow. Then for every \( t \in (0, T) \)
\[
\frac{1}{2} \partial^+_s d_{s,t}^2(x_s, y) \bigg|_{s=t} \leq E_t(y) - E_t(x_t) + (L \to \frac{K}{2}) d_{t}^2(x_t, y).
\]

2. There exists at most one \( E(V, \infty) \)-gradient flow starting in \( x' \). More precisely the following holds: Let \((x_t)\) and \((y_t)\) be two \( E(V, \infty) \)-gradient flows. Then for all \( s < t \)
\[
d_{t}(x_t, y_t) \leq e^{(3L-K)(t-s)} d_{s}(x_s, y_s).
\]
Proof. To show the first assertion note that with \( d_{t,s}(y, x_s) = d_{s,t}(x_s, y) \)

\[
\partial_s^+ d_t^2(y, x_s)_{s=t^+} := \limsup_{s \to t} \frac{d_t^2(y, x_s) - d_t^2(y, x_t)}{s - t} \\
\geq \limsup_{s \to t} \frac{e^{-2L(s-t)} d_t^2(y, x_s) - d_t^2(y, x_t)}{s - t} \\
\geq \limsup_{s \to t} \left\{ \frac{d_t^2(y, x_s) - d_t^2(y, x_t)}{s - t} + \frac{(e^{-2L(s-t)} - 1)}{s - t} d_t^2(y, x_s) \right\} \\
= \partial_s^+ d_t^2(y, x_s)|_{s=t^+} + \lim_{s \to t} \frac{(e^{-2L(s-t)} - 1)}{s - t} d_t^2(y, x_s) \\
= \partial_s^+ d_t^2(y, x_s)|_{s=t^+} - 2Ld_t^2(y, x_t),
\]

where the first inequality is due to the logarithmic Lipschitz continuity [12], and the second equality follows from the absolute continuity of \((x_t)\). The same argument holds for \( \partial_s^+ d_t^2(y, x_s)_{s=t-} := \limsup_{s \to t} \frac{d_t^2(x_s, y) - d_t^2(x_t, y)}{s - t} \) replacing \( \partial_s^+ d_t^2(y, x_s)_{s=t^+} \) by \( \partial_s^+ d_t^2(y, x_s)_{s=t^+} \), and hence from the EVI\((K, \infty)\) inequality we deduce

\[
\frac{1}{2} \partial_s^+ d_t^2(x, y)|_{s=t} \leq E_t(y) - E_t(x) + (L - \frac{K}{2})d_t^2(x, y).
\]

In order to show the second assertion, let \((x_t), (y_t)\) be two EVI\((K, \infty)\) gradient flows. Observe that from the absolute continuity of \((x_t)\) and \((y_t)\) it follows that the map \( t \to d_t^2(x_t, y_t) \) is absolutely continuous as well. This can be seen by applying triangle inequality and [12]. Hence we may write for a.e. \( t \in (0, T) \)

\[
\frac{1}{2} \frac{d}{dt} d_t^2(x_t, y_t) \leq \frac{1}{2} \limsup_{s \to t} \frac{d_t^2(x_t, y_t) - d_t^2(x_s, y_t)}{t - s} + \frac{1}{2} \limsup_{s \to t} \frac{d_t^2(x_s, y_t) - d_t^2(x_s, y_t)}{t - s} + \frac{1}{2} \limsup_{s \to t} \frac{d_t^2(x_s, y_t) - d_t^2(x_t, y_t)}{s - t}, (27)
\]

where we used an adaption of [2] Lemma 4.3.4. Applying (25) and (12) we obtain for a.e. \( t \in (0, T) \)

\[
\frac{1}{2} \frac{d}{dt} d_t^2(x_t, y_t) \leq E_t(y_t) - E_t(x_t) + (L - \frac{K}{2})d_t^2(x_t, y_t) + Ld_t^2(x_t, y_t) + E_t(x_t) - E_t(y_t) + (L - \frac{K}{2})d_t^2(x_t, y_t) = (3L - K)d_t^2(x_t, y_t).
\]

We conclude from Gronwall’s inequality for a.e. \( t > s \)

\[
d_t^2(x_t, y_t) \leq e^{(3L - 2K)(t-s)}d_s^2(x_s, y_s).
\]

From the continuity of \( t \mapsto d_t(x_t, y_t) \) we obtain that the estimate holds for every \( t > s \) and in particular we have uniqueness. 

\[ \square \]
In this general framework it is possible to produce dynamic EVI-gradient flows which are not dynamic EDI-gradient flows as we see in the next example.

**Example 5.** Let \( X = \mathbb{R} \) and \( d_t(x,y) = |x - y| \) for every \( t \in [0,T], x , y \in X \). As already seen in Example 3 the energy functional \( E_t(x) = |x - t| \) is not differentiable at \( x_t = t \) for any \( t \in [0,T] \). Hence it is not a EDI-gradient flow in the sense of Definition 5.4. But it immediately follows from

\[
\frac{1}{2} \partial_t |x_t - y|^2 = (t - y) \leq |y - t| = E_t(y) - E_t(x_t), \quad \forall y \in X,
\]

that \( (x_t) \) is a EVI(0,\( \infty \))-gradient flow.

We can exclude such behavior if we restrict ourselves to admissible functionals.

**Proposition 3.11.** Let \( E : [0,T] \times X \to \mathbb{R} \) be an admissible functional, i.e. satisfying the assumptions \( A1, A2, A3 \) and \( A4 \) from the previous section. Let \( (x_t) \) be a dynamic EVI\((K,\infty)\)-gradient flow for \( E \) such that \( (x_t) \in AC_{loc}^2([0,T]; X) \) and \( t \mapsto E_t(x_t) \) is absolutely continuous. Then it is a dynamic EDE-gradient flow as well.

**Proof.** First note that for a.e. \( t \)

\[
\frac{1}{2} \frac{\partial^+}{\partial t} d_t^2(x_t, y) \geq -|\dot{x}_t| d_t(x_t, y). \tag{28}
\]

Since \( E \) is admissible and \( t \mapsto E_t(x_t) \) is supposed to be absolutely continuous it holds for a.e. \( t \)

\[
\frac{d}{dt} E_t(x_t) = (\partial_t E_t)(x_t) + \liminf_{h \to 0} \frac{E_t(x_{t+h}) - E_t(x_t)}{h} \tag{29}
\]

\[
\geq (\partial_t E_t)(x_t) - \limsup_{h \to 0} \frac{E_t(x_t) - E_t(x_{t+h}, x_t)}{d_t(x_{t+h}, x_t)} \frac{d_t(x_{t+h}, x_t)}{h} \geq (\partial_t E_t)(x_t) - \frac{1}{2} |\nabla E_t|^2(x_t) - \frac{1}{2} |\dot{x}_t|^2.
\]

To show the converse inequality recall that \( t \mapsto d_t^2(x_t, y) \) is absolutely continuous. Hence, applying the same calculation as in (27) to the constant curve \( y_t \equiv y \), we can write for every \( t \in [0,T-h] \) and every \( y \)

\[
\frac{1}{2} d_{t+h}^2(x_{t+h}, y) - \frac{1}{2} d_t^2(x_t, y) = \frac{1}{2} \int_t^{t+h} \frac{d}{ds} d_s^2(x_s, y) ds \leq \int_t^{t+h} E_s(y) - E_s(x_s) + (2L - \frac{K}{2}) d_s^2(x_s, y) ds.
\]

We set \( y = x_t \) and find

\[
\frac{1}{2} d_{t+h}^2(x_{t+h}, x_t) \leq h \int_0^1 E_{t+hr}(x_t) - E_{t+hr}(x_{t+hr}) dr + (2L - \frac{K}{2}) \int_t^{t+h} d_s^2(x_s, x_t) dr.
\]
We additionally impose a topology
Topological assumptions we describe in the subsequent sections, cf. [2].

Again by [12] and the 2-absolute continuity of \((x_t)\) we obtain for some function \(g \in L^2_{loc}[0,T]\)

\[
\frac{1}{2} \frac{d}{dt}(x_{t+h}, x_t) \leq e^{2Lh} \left[ h \int_0^1 E_{t+hr}(x_t) - E_{t+hr}(x_{t+h}) \, dr + \left| 2L - \frac{K}{2} \right| h^2 \int_0^{t+h} g^2 \, du \right].
\]

Dividing by \(h^2\) and letting \(h \searrow 0\), dominated convergence yields

\[
\frac{1}{2} |\dot{x}_t|^2 \leq \lim_{h \searrow 0} \frac{1}{h} \int_0^1 \frac{E_t(x_t) - E_{t+hr}(x_{t+h})}{h} + \frac{E_{t+hr}(x_t) - E_t(x_t)}{h} \, dr
\]

\[
= -\frac{1}{2} \frac{d}{dt} E_t(x_t) + \frac{1}{2} (\partial_t E_t)(x_t),
\]

for a.e. \(t \in (0,T)\). Concerning the slope of \(E\) we find that using [25] and [25]

\[
|\nabla E_t(x_t)| = \lim_{y \to x_t} \frac{|E_t(x_t) - E_t(y)|^+}{d_t(x,y)}
\]

\[
\leq \lim_{y \to x_t} \frac{-\partial^+ d_t^2(x_t,y)_{s=t} + (2L - K) d_t^2(x_t,y)}{2d_t(x_t,y)}
\]

\[
\leq \lim_{y \to x_t} \frac{2|\dot{x}_t| d_t(x_t,y) + (2L - K) d_t^2(x_t,y)}{2d_t(x_t,y)}
\]

\[
\leq |\dot{x}_t|, t, \text{ for almost every } t. \quad \text{Combining (30) and (31) we conclude}
\]

\[
\frac{1}{2} \frac{d}{dt} E_t(x_t) \leq (\partial_t E_t)(x_t) - |\dot{x}_t|^2
\]

\[
= (\partial_t E_t)(x_t) - \frac{|\dot{x}_t|^2}{2} - \frac{|\nabla E_t|^2(x_t)}{2}.
\]

We obtain (18) from (20) and (32) after integrating on the interval \((0,t)\).

### 3.3 Existence of dynamic EDI-gradient flows

We are interested in the following problem.

**Problem 1.** Given a function \(E: [0,T] \times X \to (-\infty, +\infty]\), and an initial value \(\bar{x} \in \text{Dom}(E)\), find an EDI-gradient flow \((x_t)\) for \(E\).

Under suitable topological assumptions we will find a gradient flow for a certain class of energy functionals using the minimizing movement scheme, which we describe in the subsequent sections, cf. [2].

**Topological assumptions**

We additionally impose a topology \(\sigma\) on \(X\) such that \(\sigma\) is weaker than the topology induced by \((d_t)\) and \(d_t\) is sequentially \(\sigma\)-lower semicontinuous, i.e.

if \(x_n \sigma x \) and \(y_n \sigma y\), then \(\liminf_{n \to \infty} d_t(x_n, y_n) \geq d_t(x, y)\) for every \(t \in [0,T]\).

Let \(E: [0,T] \times X \to (-\infty, \infty]\) be a functional satisfying A1, A2, and A4. We will extend our assumptions by the following.

**A5** If \((x_n) \subset X\) with \(\sup_{n,m} d_t(x_n, x_m) < \infty\), then \((x_n)\) admits a \(\sigma\)-convergent subsequence.

**A3** For each \(t \in [0,T]\), \(x \mapsto E_t(x)\) is sequentially \(\sigma\)-lower semicontinuous.

18
Approximation

We fix a time step \( h > 0 \) and subdivide the interval \([0, T]\) into the partition
\[
\mathcal{P}_h := \{t_0 = 0 < t_1 < \cdots < t_{N-1} < T \leq t_N\}, \quad t_n = nh, N \in \mathbb{N}.
\]
For \( 0 \leq t \leq T \) we define the piecewise constant interpolants \( \overline{h}(t) \) and \( h(t) \) associated with the partition \( \mathcal{P}_h \) in the following way;
\[
\overline{h}(0) = 0 = h(0), \quad \text{and for } t \in (t_{n-1}, t_n] \quad \overline{h}(t) = t_n, \quad h(t) = t_{n-1}.
\]
The definition implies that \( \overline{h}(t) \searrow t \) and \( h(t) \nearrow t \) if \( h \searrow 0 \).

For a given initial value \( \bar{x} \) we recursively define a sequence \( (x_n^h) \) of minimizers by
\[
x_0^h := \bar{x}, \quad x_n^h := \arg \min_x \left\{ E_{t_n}(x) + \frac{1}{2h} d_{x_{n-1}}^2(x, x_n^h) \right\} \tag{34}
\]

**Proposition 3.12.** For every \( \bar{x} \in \text{Dom}(E) \) and \( h > 0 \) there exists a solution to the minimization problem \( (34) \).

**Proof.** Existence follows by the direct method of calculus. Define
\[
\phi(h, \bar{x}, t; \cdot) := E_t(\cdot) + \frac{1}{2h} d_t^2(\bar{x}, \cdot).
\]
Since \( E \) is uniformly bounded from below we may take a minimizing sequence \( (x_n)_{n \in \mathbb{N}} \) such that \( d_t^2(x_n, \bar{x}) \) remains bounded uniformly in \( n \). From the triangle inequality we deduce that \( \sup_{n,m} d_t(x_n, x_m) < \infty \). Hence \( A5 \) guarantees existence of a \( \sigma \)-convergent subsequence \( x_{n_k} \). The weak limit point \( x \in \text{Dom}(E) \) is a minimizer of \( \phi(h, \bar{x}, t; \cdot) \), which is due to the \( \sigma \)-lower semicontinuity of the distance and the functional. \( \square \)

**Definition 3.13.** Fix \( h > 0 \) and let \( s \in [0, T - h] \). For \( 0 < r < T - s \) define
\[
J_{s,r}(y) := \min_x \left\{ E_{s+r}(x) + \frac{1}{2r} d_{s+h}^2(x, y) \right\}, \tag{35}
\]
\[
A_{s,r}(y) := \arg \min_x \left\{ E_{s+r}(x) + \frac{1}{2r} d_{s+h}^2(x, y) \right\}. \tag{36}
\]

**Lemma 3.14.** For \( x_r \in A_{s,r}(y) \) we have
\[
|\nabla_{s+h} E_{s+r}(x_r) | \leq \frac{1}{r} d_{s+h}(x_r, y)
\]
and for \( 0 < r_1 < r_2 < T - s \)
\[
d_{s+h}^2(x_{r_1}, y) \leq d_{s+h}^2(x_{r_2}, y) + 4r_1 r_2 L^* \tag{37}
\]
where \( L^* \) is the Lipschitz constant in assumption \( A4 \).

**Proof.** By optimality of \( x_r \) we have for every \( x \in X \)
\[
\frac{E_{s+r}(x_r) - E_{s+r}(x)}{d_{s+h}(x_r, x)} \leq \frac{d_{s+h}^2(x, y) - d_{s+h}^2(x_r, y)}{2rd_{s+h}(x_r, x)} \leq \frac{(d_{s+h}(x, y) - d_{s+h}(x_r, y))(d_{s+h}(x, y) + d_{s+h}(x_r, y))}{2rd_{s+h}(x_r, x)} \leq \frac{(d_{s+h}(x, y) + d_{s+h}(x_r, y))}{2r}.
\]

19
Taking the limsup as $x \to x_r$ we get the first assertion. To show the second assertion note that on the one hand we have
\[
E_{s+r_1}(x_{r_1}) + \frac{1}{2r_1} d_{s+h}^2(x_{r_1}, y) \leq E_{s+r_1}(x_{r_2}) + \frac{1}{2r_1} d_{s+h}^2(x_{r_2}, y),
\]
and on the other
\[
E_{s+r_2}(x_{r_2}) + \frac{1}{2r_2} d_{s+h}^2(x_{r_2}, y) \leq E_{s+r_2}(x_{r_1}) + \frac{1}{2r_2} d_{s+h}^2(x_{r_1}, y).
\]
Adding these two inequalities, using the Lipschitz property of $r \mapsto E_t(x)$ and dividing by $\frac{1}{2r_1} - \frac{1}{2r_2}$ yields $37$.

**Lemma 3.15.** The map $r \mapsto J_{s,r}(y)$ is locally Lipschitz and for almost every $r \in (0, T - s)$ we have for $x_r \in A_{s,r}(y)$
\[
\frac{d}{dr} J_{s,r}(y) = -\frac{1}{2r^2} d_{s+h}^2(x_r, y) + (\partial_r E_{s+r})(x_r).
\]
**Proof.** Fix $0 < r_1 < r_2 < T - s$. Then
\[
J_{s,r_2}(y) - J_{s,r_1}(y) = E_{s+r_2}(x_{r_2}) - E_{s+r_1}(x_{r_1})
+ \frac{1}{2r_2} d_{s+h}^2(x_{r_2}, y) - \frac{1}{2r_1} d_{s+h}^2(x_{r_1}, y)
\leq E_{s+r_2}(x_{r_1}) - E_{s+r_1}(x_{r_1}) + \frac{r_1 - r_2}{2r_2 r_1} d_{s+h}^2(x_{r_1}, y)
\leq L^*(r_2 - r_1) - \frac{r_2 - r_1}{2r_2 r_1} d_{s+h}^2(x_{r_1}, y),
\]
where $L^*$ denotes the Lipschitz constant from A4. Conversely, changing the roles of $x_{r_1}$ and $x_{r_2}$, we obtain
\[
J_{s,r_2}(y) - J_{s,r_1}(y) \geq -L^*(r_2 - r_1) - \frac{r_2 - r_1}{2r_2 r_1} d_{s+h}^2(x_{r_2}, y).
\]
Combining these two inequalities yields
\[
|J_{s,r_2}(y) - J_{s,r_1}(y)| \leq L^* |r_2 - r_1| + \frac{|r_2 - r_1|}{2r_2 r_1} d_{s+h}^2(x_{r_2}, y),
\]
which means $r \mapsto J_{s,r}(y)$ is locally Lipschitz. Dividing by $r_2 - r_1$ and letting $r_1 \to r_2$ in $38$ yields on the one hand for the left derivative
\[
\frac{d^-}{dr} J_{s,r}(y) \leq -\frac{1}{2r^2} d_{s+h}^2(x_r, y) + (\partial_r E_{s+r})(x_r),
\]
for every differentiability point $r$ of $r \mapsto E_{t+r}$. On the other hand we obtain similarly for the right derivative
\[
\frac{d^+}{dr} J_{s,r}(y) \geq -\frac{1}{2r^2} d_{s+h}^2(x_r, y) + (\partial_r E_{s+r})(x_r),
\]
for every differentiability point of $r \mapsto E_{t+r}$. By local Lipschitz continuity we have for a.e. $0 < r < T - s$
\[
\frac{d}{dr} J_{s,r}(y) = -\frac{1}{2r^2} d_{s+h}^2(x_r, y) + (\partial_r E_{s+r})(x_r).
\]
Lemma 3.16. For $s \in [0,T]$ and $0 < r_1 < r_2 < T - s$

\[ E_s(y) \geq J_{s,r_1}(y) - Cr_1 \geq J_{s,r_2}(y) - Cr_2 \tag{40} \]

\[ \lim_{r \to 0} d_{s+h}(y,x_r) = 0 \text{ if } y \in \text{Dom}(E). \tag{41} \]

In particular \( \lim_{r \to 0} J_{s,r}(y) = E_s(y) \).

**Proof.** The first inequality in (40) directly follows from (34) defined by the minimization problem

\[ \text{and minimizing over all } x \in \text{Dom}(E). \]

The second one follows by Corollary 3.17. For every $y \in \text{Dom}(E)$

\[ 0 \leq d^2_{s+h}(y,x_r) \leq -2rE_{s+r}(x_r) + d^2_{s+h}(y,x) + 2rE_{s+r}(x) \leq -2r \inf \{E + d^2_{s+h}(y,x) + 2rE_{s+r}(x)\}. \]

Passing to the limit \( r \to 0 \)

\[ \lim_{r \to 0} d^2_{s+h}(x_r,y) \leq d^2_{s+h}(x,y) \text{ for every } x \in \text{Dom}(E). \]

Since $y \in \text{Dom}(E)$ we conclude (41). To check the last one we combine (40) with the lower semicontinuity of $x \mapsto E_t(x)$,

\[ E_t(y) \geq \liminf_{r \to 0} J_{t,r}(y) \geq \liminf_{r \to 0} E_{t+r}(x_r) \geq E_t(y). \]

\[ \square \]

**Corollary 3.17.** For every $0 < r_0 < T - s$ we have

\[ E_{s+r_0}(x_{r_0}) + \frac{1}{2r_0}d^2_{s+h}(x_{r_0},y) \]

\[ = E_s(y) - \int_0^{r_0} \frac{1}{2r^2}d^2_{s+h}(x_r,y)dr + \int_0^{r_0} (\partial_r E_{s+r})(x_r)dr. \tag{42} \]

**Proof.** Integrate (42) from 0 to $r_0$ and use that $\lim_{r \to 0} J_{s,r}(y) = E_s(y)$. \[ \square \]

In the following we introduce dynamic counterparts for the variational interpolation, the discrete speed and the discrete slope, cf. [2] [21].

**Definition 3.18.** Let $x \in \text{Dom}(E)$ be the initial value and $x^n_h$ be a sequence defined by the minimization problem (55). A discrete solution is a curve $t \mapsto \dot{x}^n_t$ defined by

\[ \dot{x}^n_t = x^n_{t_n}, \text{ for } t \in (t_{n-1},t_n], \]
and \( \tilde{x}_0^h = \tilde{x} \).

A variational interpolation is a map \( t \to \tilde{x}_t^h \) defined by

\[
\tilde{x}_t^h = \arg \min \left\{ E_t(x) + \frac{1}{2r} d_t^2(x, x_{n-1}^{h}) \right\},
\]

for \( t = t_{n-1} + r \in (t_{n-1}, t_n] \),

and \( \tilde{x}_0^h = \tilde{x} \).

We define the discrete speed \( \text{Dis}^h \colon [0, T] \to [0, \infty) \) and the discrete slope \( \text{Dis}^h \colon [0, T] \to [0, \infty) \) in the following way

\[
\text{Dis}^h_r = \frac{1}{h} d_t^2(x_{t_{n-1}}^h, x_{t_n}^h), \quad r \in (t_{n-1}, t_n],
\]

\[
\text{Dis}^h_r = \frac{1}{r - t_{n-1}} d_t^2(x_{t_{n-1}}^h, x_{t_r}^h), \quad r \in (t_{n-1}, t_n].
\]

Note that \( \tilde{x}_{t_n}^h = x_{t_n}^h = \tilde{x}_{t_n}^h \).

**Proposition 3.19.** We have for \( 0 \leq s \leq t \leq T \)

\[
E_{n(t)}(\tilde{x}_t^h) + \frac{1}{2} \int_{[n(s)]}^{n(t)} (\text{Dis}_r^h)^2 dr + \frac{1}{2} \int_{[n(s)]}^{n(t)} (\text{Dis}_r^h)^2 dr = E_{n(s)}(\tilde{x}_s^h) + \int_{[n(s)]}^{n(t)} (\partial_r E_r)(\tilde{x}_r^h) dr.
\]

**Proof.** Let \( t \in (t_{n-1}, t_n] \). We want to apply equation (43) with \( s = t_{n-1}, r_0 = t - s, y = x_{t_{n-1}}^h \). Then with \( x_{r_0} = \tilde{x}_t^h \) and \( x_r = \tilde{x}_{t_{n-1}+r}^h \) we find

\[
E_t(\tilde{x}_t^h) + \frac{1}{2(t - t_{n-1})} d_t^2(\tilde{x}_t^h, x_{t_{n-1}}^h) + \int_{t_{n-1}}^{t} \frac{1}{2r(t - t_{n-1})^2} d_t^2(x_{t_{n-1}}^h, \tilde{x}_r^h) dr
= E_{t_{n-1}}(x_{t_{n-1}}^h) + \int_{t_{n-1}}^{t} (\partial_r E_r)(\tilde{x}_r^h) dr.
\]

For \( t = t_n \) we obtain

\[
E_{t_n}(\tilde{x}_{t_n}^h) + \frac{1}{2h^2} \int_{t_{n-1}}^{t_n} d_t^2(\tilde{x}_{t_n}^h, x_{t_{n-1}}^h) dr + \int_{t_{n-1}}^{t_n} \frac{1}{2r(t - t_{n-1})^2} d_t^2(x_{t_{n-1}}^h, \tilde{x}_r^h) dr
= E_{t_{n-1}}(\tilde{x}_{t_{n-1}}^h) + \int_{t_{n-1}}^{t_n} (\partial_r E_r)(\tilde{x}_r^h) dr.
\]

Summing up from \( n+1 \) to \( m \) yields

\[
E_{t_m}(\tilde{x}_{t_m}^h) + \frac{1}{2h^2} \sum_{j=n+1}^{m} \int_{t_{j-1}}^{t_j} d_t^2(\tilde{x}_{t_j}^h, x_{t_{j-1}}^h) dr
+ \sum_{j=n+1}^{m} \int_{t_{j-1}}^{t_j} \frac{1}{2(r - t_{j-1})^2} d_t^2(x_{t_{j-1}}^h, \tilde{x}_r^h) dr = E_{t_{n}}(\tilde{x}_{t_{n}}^h) + \int_{t_{n}}^{t_m} (\partial_r E_r)(\tilde{x}_r^h) dr.
\]

22
Now plugging in the definitions of the discrete slope and the discrete speed respectively
\[
E_{tm}(\bar{x}_{tn}^h) + \frac{1}{2} \int_{tn}^{tm} (Dsp_r^h)^2 dr + \frac{1}{2} \int_{tn}^{tm} (Dsl_r^h)^2 dr
= E_{tn}(\bar{x}_{tn}^h) + \int_{tn}^{tm} (\partial E_r)(\bar{x}_{tn}^h) dr,
\]
which shows (43).

Remark 3.20. Alternatively, for \( t \in (t_{n-1}, t_n] \) we can write
\[
E_t(\bar{x}_t^h) + \frac{1}{2(t - t_{n-1})} d_r^2(\bar{x}_t^h, \bar{x}_{t_{n-1}}^h) + \frac{1}{2} \int_{t_{n-1}}^{t} (Dsp_r^h)^2 dr + \frac{1}{2} \int_{0}^{t} (Dsl_r^h)^2 dr
= E_0(\bar{x}) + \int_{0}^{t} (\partial E_r)(\bar{x}_r^h) dr.
\]
The following proposition provides essential a priori bounds.

Proposition 3.21. There exist constants \( C_1, C_2, C_3 \) such that for all \( 0 \leq t, mh \leq T \)
\[
E_t(\bar{x}_t^h) \leq C_1,
\]
\[
\frac{1}{2h} \sum_{n=1}^{m} d_r^2(x_{tn}, \bar{x}_{tn-1}) \leq C_2,
\]
\[
d_r^2(\bar{x}_t^h, \bar{x}_{tn}^h) \leq C_3 h, \text{ for some fixed } t^*.
\]

Proof. From Remark 3.20 we deduce
\[
E_t(\bar{x}_t^h) \leq E_0(\bar{x}) + L^* T,
\]
which shows (45).

We drop the nonnegative slope term in equation (44) to obtain
\[
\frac{1}{2h} d_r^2(x_{tn}, \bar{x}_{tn-1}) \leq E_{tn-1}(\bar{x}_{tn-1}) - E_{tn}(\bar{x}_{tn}) + \int_{tn-1}^{tn} (\partial E_r)(\bar{x}_{tn}^h) dr.
\]
Summing up to \( m \) and applying the Lipschitz property of \( t \rightarrow E_t \)
\[
\frac{1}{2h} \sum_{n=1}^{m} d_r^2(x_{tn}, \bar{x}_{tn-1}) \leq E_{tn-1}(x_{tn-1}) - E_{tn}(x_{tn}) + \int_{0}^{t} (\partial E_r)(\bar{x}_r^h) dr
\leq E_{tn}(\bar{x}_{tn}) - E_{tn}(\bar{x}_t^h) + TL^*,
\]
we obtain on the one hand
\[
E_{tm}(\bar{x}_{tm}^h) \leq E_{t_0}(\bar{x}_{t_0}) + TL^*,
\]
and since \( \inf E(x) > -\infty \)
\[
\frac{1}{2h} \sum_{n=1}^{m} d_r^2(x_{tn}, \bar{x}_{tn-1}) \leq C_2.
\]
Further Proof. We want to apply a refined version of Arzelà-Ascoli \cite[Proposition 3.3.1]{2} to the family $(\tilde{x}^h)_{n \geq 0}$, relabeled, and a limit curve $x$. Proposition 3.3.1 in \cite{2} and (47) imply that there exists a further subs equence, where the third inequality is a consequence of \cite{37}. Applying \cite{40} and \cite{12} we conclude \cite{17}.

**Proposition 3.22.** There exist a subsequence $h_n$ with $\lim_n h_n = 0$, a curve $(x_t) \subset AC^2([0, T]; X)$ and a function $A \in L^2(0, T)$ such that

\[ \dot{x}^h_t \preceq x_t, \quad x^h_t \xrightarrow{\sigma} x_t \text{ for all } t, \]

and $|Dsp^h| \rightarrow A$ weakly in $L^2(0, T)$.

Further $|\dot{x}|_t \leq A(t)$ holds almost everywhere.

**Proof.** We want to apply a refined version of Arzelà-Ascoli \cite[Proposition 3.3.1]{2} to the family $(\tilde{x}^h)_{h > 0}$. Owing to the estimates \cite{12} and \cite{40} we have

\[ d_{t^*}(\tilde{x}^h_t, \hat{x}) \leq tC_2e^{C_t}, \]

and together with A5 this yields that the curves $\tilde{x}^h : [0, T] \rightarrow X$ take values in a $\sigma$-sequentially compact set. From the estimate \cite{40} we further deduce

\[ \int_t^s |Dsp^h|^2 |dr \leq \sum_{j=t_n}^{t_{n-1}} \frac{1}{h} d_{t_j}(\tilde{x}^h_j, \dot{x}^h_j) \leq 2C_2, \]

for $\tilde{h}(s) = t_m, \tilde{h}(t) = t_n$. Applying the Banach Alaoglu Theorem we can extract a subsequence $h_n$ and a function $A \in L^2(0, T)$ such that $|Dsp^h| \rightarrow A$ weakly in $L^2([0, T])$. For fixed $t^*$ and $s < t$ we deduce from the log-Lipschitz property \cite{12}

\[ d_{t^*}(\tilde{x}^h_t, \dot{x}^h_s) \leq \int_{\tilde{h}(s)}^{\tilde{h}(t)} \frac{1}{h} d_{t^*}(\dot{x}^h_{t^*}, \dot{x}^h_{t^*}) |dr, \]

and hence

\[ \lim_{n \rightarrow \infty} d_{t^*}(x^h_t, x^h_s) \leq \int_t^s A(r)e^{L[r-t^*]} |dr. \]

Proposition 3.3.1 in \cite{2} and \cite{17} imply that there exists a further subsequence, not relabeled, and a limit curve $x : [0, T] \rightarrow X$ such that

\[ \tilde{x}^h_t \xrightarrow{\sigma} x_t, \quad \tilde{x}^h_s \xrightarrow{\sigma} x_t \quad \forall t \in [0, T]. \]

This curve is absolutely continuous since

\[ d_{t^*}(x_t, x_s) \leq \liminf_{n \rightarrow \infty} d_{t^*}(\tilde{x}^h_t, \tilde{x}^h_s) \leq \int_s^t A(r)e^{L[r-t^*]} |dr, \]

24
In particular if we take \(t^* = t\) in the argumentation above the Lebesgue differentiation theorem implies that

\[
|\dot{x}| \leq \limsup_{s \nearrow t} \frac{1}{t-s} \int_{s}^{t} A(r) e^{E_{t}(r-t)} dr \leq A(t)
\]

holds true for almost every \(t\).

**Proposition 3.23.** Suppose additionally to our standing assumptions A1, A2, A3*, A4 and A5 that

- If \(x_n \rightharpoonup^s x\) as \(n \to \infty\) then
  \[
  \limsup_{n \to \infty} \partial_t E_t(x_n) \leq \partial_t E_t(x),
  \]
  (48)

- if \(t_n \to t\) and \(x_n \rightharpoonup^s x\), then
  \[
  |\nabla_t E_t|^2(x) \leq \liminf_{n \to \infty} |\nabla_{t_n} E_{t_n}|^2(x_n).
  \]

Then every limit curve \((x_t)\) from Proposition 3.22 satisfies the EDI formula

\[
E_t(x_t) + \frac{1}{2} \int_{0}^{t} |\dot{x}|^2 dr + \frac{1}{2} \int_{0}^{t} |\nabla_r E_r|^2(x_r) dr \leq E_0(\bar{x}) + \int_{0}^{t} (\partial_r E_r)(x_r) dr,
\]

(49)

for every \(t \in [0,T]\).

**Proof.** Recall that Proposition 3.19 states for \(s = 0\)

\[
E_{\bar{n}(t)}(\bar{x}_t^{h_n}) + \frac{1}{2} \int_{0}^{\bar{n}(t)} D_{s}\bar{r}^{h_n})^2 dr + \frac{1}{2} \int_{0}^{\bar{n}(t)} (D_{s}\bar{r}^{h_n})^2 dr
= E_0(\bar{x}) + \int_{0}^{\bar{n}(t)} (\partial_r E_r)(\bar{x}_r^{h_n}) dr.
\]

Since both \(\bar{x}_n\), \(\bar{x}_n^{h_n}\) \(\sigma\)-converges to \(x_t\) for every \(t\), \(x \mapsto E_t(x)\) is \(\sigma\)-lower semicontinuous, \(x \mapsto \partial_1 E_t(x)\) is \(\sigma\)-upper semicontinuous and \(t \mapsto E_t(x)\) is Lipschitz continuous uniformly in \(x\), we know

\[
\liminf_{n \to \infty} E_{\bar{n}(t)}(\bar{x}_t^{h_n}) \geq E_t(x_t),
\]

and

\[
\int_{0}^{t} (\partial_r E_r)(x_r) dr \geq \int_{0}^{t} \limsup_{n \to \infty} (\partial_r E_r)(\bar{x}_r^{h_n}) dr \geq \liminf_{n \to \infty} \int_{0}^{t} (\partial_r E_r)(\bar{x}_r^{h_n}) dr,
\]

where the last inequality follows from Fatou’s Lemma. From Proposition 3.22 and Lemma 3.14 we deduce

\[
\int_{0}^{t} |\dot{x}|^2 dr \leq \int_{0}^{t} A(r)^2 dr \leq \liminf_{n \to \infty} \int_{0}^{t} (D_{s}\bar{r}^{h_n})^2 dr,
\]

25
and
\[ \int_0^t |\nabla_x E_r|^2(x_r) dr \leq \lim \inf \int_0^t |\nabla_x E_r|^2(\tilde{x}_r^h) dr \leq \lim \inf \int_0^t (Ds_l^h)^2 dr. \]

Combining these inequalities with (43) we conclude
\[
E_t(x_t) + \frac{1}{2} \int_0^t |\dot{x}|^2 dr + \frac{1}{2} \int_0^t |\nabla_x E_r|^2(x_r) dr \\
\leq \lim \inf \left[ E_{\pi_n(t)}(\tilde{x}_r^h) + \frac{1}{2} \int_0^{\pi_n(t)} (Ds_r^h)^2 dr + \frac{1}{2} \int_0^{\pi_n(t)} (Ds_l^h)^2 dr \right] \\
\leq \lim \inf \left[ E_0(\bar{x}) + \int_0^t (\partial_r E_r)(\tilde{x}_r^h) dr \right] \\
\leq E_0(\bar{x}) + \int_0^t (\partial_r E_r)(x_r) dr,
\]
which is the assertion.

\[ \square \]

4 Dynamic gradient flow of the entropy

In this section we want to study gradient flows for the Boltzmann entropy on probability space, where the metric of the space and the reference measure of the entropy varies in time. To show existence we apply the results from Section 3.3. We then go on to show also uniqueness.

Let \( X \) be a topological space equipped with a family of complete separable geodesic metrics \((d_t)_{t \in [0,T]}\) satisfying (12) and a Borel probability measure \( m \).

We define \( \mathcal{P}(X) \) to be the space of Borel probability measures on \( X \) and we denote the subspace of probability measures absolutely continuous to the measure \( m \) by \( \mathcal{P}^{ac}(X) \). Further let \( \mathcal{P}_2(X) \) be the space of probability measures with finite second moments on \( X \)

\[ \mathcal{P}_2(X) := \left\{ \mu \in \mathcal{P}(X) \mid \int d_t^2(x,x_0) d\mu(x) < \infty \right\}. \]

We say that a sequence \( \mu_n \subset \mathcal{P}(X) \) converges weakly to \( \mu \) if \( \lim \int_X f d\mu_n = \int_X f d\mu \) for every \( f \in C^0_b(X) \). We say that a sequence \( \rho_n \subset L^1(X,m) \) converges weakly to \( \rho \) if \( \lim \int_X f \rho_n dm = \int_X f dm \) for every \( f \in L^\infty(X,m) \). Note that if \( \rho_n \) converges weakly to \( \rho \) in \( L^1(X,m) \) then \( \mu_n = \rho_n m \) converges weakly to \( \mu = \rho m \) in \( \mathcal{P}(X) \).

4.1 Time-dependent Kantorovich metrics

For every metric \( d_t \) we define the \( L^2 \)-Kantorovich distance \( W_t \) on the space \( \mathcal{P}_2(X) \):
\[ W_t(\mu,\nu) = \inf \{ C_t(\gamma) : \pi_{\#}^1 \gamma = \mu, \pi_{\#}^2 \gamma = \nu \}^{1/2}, \]

26
where $C_t(\gamma)$ is the cost of the plan $\gamma \in \mathcal{P}(X \times X)$

$$C_t(\gamma) = \int d_t^2(x, y) d\gamma(x, y),$$

and $\pi^1, \pi^2$ denote the first and second marginals of $\gamma$.

For each $t \in [0, T]$, $(\mathcal{P}_2(X), W_t)$ is a geodesic Polish space, see e.g. [26, 1]. It is well-known that convergence in the $L^2$-Kantorovich distance $W_t$ implies weak convergence in $\mathcal{P}(X)$ and that $W_t$ is lower semicontinuous on $\mathcal{P}(X)$ (cf. [26, Theorem 6.8] and [26, Remark 6.10]). The bound (12) is equivalent to

$$|\log W_t(\mu, \nu)/W_s(\mu, \nu)| \leq L|t - s|,$$

(50)

for all $s, t$ and all probability measures on $X$, see Lemma 2.1 in [24].

The convexity of the squared metric speed is crucial for showing uniqueness of the gradient flow. More precisely we have the following result [11, Lemma 14].

**Lemma 4.1.** Let $(\mu^1_t), (\mu^2_t) \in AC^2([0, T]; \mathcal{P}_2(X))$ be two absolutely continuous curves. Define $\mu^{1,2}_t = (\mu^1_t + \mu^2_t)/2$. Then $(\mu^{1,2}_t)$ is absolutely continuous and the following bound on its metric derivative holds

$$|\dot{\mu}^{1,2}_t| \leq |\dot{\mu}^1_t|^2 + |\dot{\mu}^2_t|^2.$$

**Proof.** Fix $s, t \in [0, T]$. Pick optimal plans $\gamma^1, \gamma^2$, which minimize $W_t(\mu^1_t, \mu^1_s)$ and $W_t(\mu^2_t, \mu^2_s)$ respectively. Then the plan $(\gamma^1 + \gamma^2)/2$ has marginals $\mu^{1,2}_t$ and $\mu^{1,2}_s$ and therefore it holds

$$W^2_t(\mu^{1,2}_t, \mu^{1,2}_s) \leq \int d_t^2(x, y) d\gamma(x, y) \frac{1}{2} \left( |\gamma^1 - \gamma^2| \right)^2(x, y)
\begin{align*}
= \frac{1}{2} \int d_t^2(x, y) d\gamma^1(x, y) + \frac{1}{2} \int d_t^2(x, y) d\gamma^2(x, y) \\
= \frac{1}{2} W^2_t(\mu^1_t, \mu^1_s) + \frac{1}{2} W^2_t(\mu^2_t, \mu^2_s).
\end{align*}
$$

Thus the curve $(\mu^{1,2}_t)$ is absolutely continuous. Dividing by $(s - t)^2$ and taking the superior limit as $s$ goes to $t$ we get for its speed

$$|\dot{\mu}^{1,2}_t| \leq \frac{|\dot{\mu}^1_t|^2 + |\dot{\mu}^2_t|^2}{2}.$$

\[ \square \]

### 4.2 Time-dependent Boltzmann entropy

We consider a family of measures $(m_t)_{t \in [0, T]}$ on $X$. We suppose that for every $t \in [0, T]$ there exists a function $f_t \in L^\infty(X, m)$ such that $m_t = e^{-f_t}m$. Moreover let us always assume that there exists a constant $L^*$ such that

$$|f_t(x) - f_s(x)| \leq L^*|t - s|$$

(51)

for all $s, t$ and all $x$. 27
We denote by $S_t$ the relative Boltzmann entropy with respect to $m_t$,

$$S_t: [0, T] \times \mathcal{P}_2(X) \to [-\infty, \infty],$$

$$(t, \mu) \mapsto S_t(\mu) = \text{Ent}(\mu | m_t) = \int \rho \log \rho \, dm_t,$$

where $\rho = d\mu / dm_t$ provided that $\mu \ll m_t$. Otherwise we set $S_t(\mu) = \infty$. It follows directly from the representation of the measures $m_t$ that

$$S_t(\mu) = \text{Ent}(\mu) + \int f_t(x) \, d\mu(x),$$

where $\text{Ent}(\mu) = \text{Ent}(\mu | m)$. In the next lemma we list the crucial properties of the relative entropy functional.

**Lemma 4.2.** The entropy $S_t: [0, T] \times \mathcal{P}_2(X) \to [-\infty, \infty]$ satisfies $A_1, A_2, A_3^*$ and $A_4$, i.e.

1. The domain $\text{Dom}(S_t)$ is time-independent.
2. $S_t(\mu)$ is uniformly bounded from below.
3. For each $t \in [0, T]$, $\mu \mapsto S_t(\mu)$ is lower semicontinuous with respect to weak convergence over probability space.
4. For every $\mu \in \text{Dom}(S)$ the map $t \mapsto S_t(\mu)$ is Lipschitz continuous with Lipschitz constant $L^*$ and for the derivative it holds

$$\partial_t S_t(\mu) = \int_X \partial_t f_t(x) \, d\mu(x) \text{ for a.e. } t \in [0, T].$$

Moreover the set of differentiability points of $t \mapsto S_t(\mu)$ can be chosen independent of $\mu$.

**Proof.** The domain is time-independent by virtue of (51). Since $m(X) = 1$ we can estimate $\text{Ent}(\mu) \geq 0$ and hence for $\mu \in \mathcal{P}^{ac}(X)$

$$S_t(\mu) \geq \int f_t(x) \, d\mu(x) \geq -||f_t||_{L^\infty} \geq -||f_\mu||_{L^\infty} - L^* T.$$

If $\mu \notin \mathcal{P}^{ac}(X)$ we know that $S_t(\mu) = \infty$ and we conclude $\inf_{t, \mu} S_t(\mu) > -\infty$.

For every $t$ the measure $m_t(X)$ is finite and thus $\mu \mapsto S_t(\mu)$ is lower semicontinuous with respect to weak convergence (Lemma 4.1 in [23]).

Fix $\mu \in \text{Dom}(S)$. The Lipschitz continuity of $t \mapsto f_t(x)$ ensures $|\partial_t S_t(\mu)| \leq C$. It is clear that for every $x \in X$ the map $t \mapsto f_t(x)$ is differentiable for a.e. $t \in [0, T]$. Hence the integral $\int_X \int_{t_1}^{t_2} |\partial_t f_t(x)| \, dtd\mu(x)$ exists and the Fubini-Tonelli theorem states

$$\int_X \int_{t_1}^{t_2} \partial_t f_t(x) \, dtd\mu(x) = \int_{t_1}^{t_2} \int_X \partial_t f_t(x) \, d\mu(x) \, dt.$$

The Fubini theorem again yields that for a.e. $t$ the map $x \mapsto \partial_t f_t(x)$ is $\mu$-integrable and so for a.e. $t$ the integral

$$\int_X \partial_t f_t(x) \, d\mu(x)$$
exists. Take a differentiability point \( t \) of \( f_t(x) \). Then for \( \mu \)-a.e. \( x \in X \)

\[
\lim_{h \to 0} \frac{1}{h}(f_{t+h} - f_t)(x) = \partial_t f_t(x), \quad \text{and} \quad \frac{1}{h}|(f_{t+h} - f_t)(x)| \leq L^*.
\]

Hence we conclude that for a.e. \( t \in [0, T] \)

\[
\lim_{h \to 0} \frac{1}{h}[S_{t+h}(\mu) - S_t(\mu)] = \lim_{h \to 0} \int \frac{1}{h}[f_{t+h}(x) - f_t(x)]d\mu(x) = \int \partial_t f_t(x)d\mu(x),
\]

where the last equality is due to the dominated convergence theorem. Finally, for \( \mu \ll m \), the inclusions

\[
\left\{ t \in [0, T] \right\} \mid \lim \frac{1}{h}[S_{t+h}(\mu) - S_t(\mu)] \text{ exists}
\]

\[
\subset \left\{ t \in [0, T] \right\} \mid \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)] \text{ exists for } \mu \text{ a.e. } x
\]

\[
\subset \left\{ t \in [0, T] \right\} \mid \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)] \text{ exists for } m \text{ a.e. } x
\]

show that the set of differentiability points of \( t \mapsto S_t(\mu) \) does not depend on \( \mu \), since the complement \( \left\{ t \in [0, T] \right\} \mid \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)] \text{ exists for } m \text{ a.e. } x \) is negligible.

Since we want to apply the results from Section 3.3 we still need to check the assumptions in Proposition 3.23. It has been shown in [11] that if the Ricci curvature of \((X, d_t, m_t)\) is bounded from below by \( K \in \mathbb{R} \) the squared slope of the entropy is lower semicontinuous. We briefly recall the arguments.

**Definition 4.3.** The set \( GP \subset \mathcal{P}(X^2) \) is the set of plans \( \gamma \) such that

1. the marginals \( \pi_{|\#}^X \gamma \), \( i = 1, 2 \) are absolutely continuous with densities bounded away from 0 and \( \infty \),
2. \( \sup_{(x,y) \in \text{supp}(\gamma)} d_h(x,y) < \infty \) for some \( t \in [0, T] \), and thus for any.

Given \( \gamma \in GP \) and \( \mu \in \mathcal{P}_2^a(X) \), we define the plan \( \gamma_\mu \in \mathcal{P}(X^2) \) and the measure \( \nu_{\gamma, \mu} \in (X) \) as

\[
d_{\gamma_\mu}(x, y) = \frac{d\mu(x)}{d\pi_{|\#}^X \gamma(x)}d\gamma(x, y), \quad \nu_{\gamma, \mu} = \pi_{|\#}^X \gamma_\mu.
\]

Note that since \( \gamma_\mu \ll \gamma \), we have \( \nu_{\gamma, \mu} \ll m \) with density

\[
g_{\gamma, \mu}(y) = \frac{d\pi_{|\#}^X \gamma(y)}{dm(y)} \int \frac{d\mu(x)}{d\pi_{|\#}^X \gamma(x)}d\gamma_y(x),
\]

where \( (\gamma_y)_y \subset \mathcal{P}(X) \) is the disintegration of \( \gamma \) with respect to its second marginal.

Observe that from 2. of the definition of the set \( GP \) we have that the cost \( C_t(\gamma) \) of a plan \( \gamma \in GP \) is always finite and \( \nu_{\gamma, \mu} \in \mathcal{P}_2(X) \) since \( \mu \in \mathcal{P}_2(X) \).

The next Proposition gives an alternative representation formula for the slope in terms of good plans, cf. [11] Theorem 12.29
Proposition 4.4. For every $t \in [0,T]$ and every $\mu \in \text{Dom}(S)$ it holds

$$
\sup_{\nu \in \mathcal{P}_2(X), \nu \neq \mu} \frac{(S_t(\mu) - S_t(\nu) - \frac{K}{2} W_t^2(\mu, \nu))^+}{W_t(\mu, \nu)} = \sup_{\gamma \in \text{GP}} \frac{(S_t(\mu) - S_t(\nu_{\gamma, \mu}) - \frac{K}{2} C_t(\gamma_{\mu}))^+}{\sqrt{C_t(\gamma_{\mu})}},
$$

where the value of the second expression is taken by definition as 0 if $C_t(\gamma_{\mu}) = 0$.

Proof. We start with proving $\geq$. For this fix a plan $\gamma \in \text{GP}$ such that $\nu_{\gamma, \mu} \neq \mu$. From $C_t(\gamma_{\mu}) \geq W_t^2(\mu, \nu_{\gamma, \mu}) > 0$ we obtain

$$
(S_t(\mu) - S_t(\nu_{\gamma, \mu}) - \frac{K}{2} W_t^2(\mu, \nu_{\gamma, \mu}))^+ \geq (S_t(\mu) - S_t(\nu_{\gamma, \mu}) - \frac{K}{2} C_t(\gamma_{\mu}))^+ \frac{C_t(\gamma_{\mu})}{\sqrt{C_t(\gamma_{\mu})}}.
$$

To show the reverse inequality take $\nu \in \mathcal{P}_2^c(X)$ different from $\mu$. Lemma 10 in [11] provides a sequence $(\gamma^n) \subset \text{GP}$ such that $S_t(\nu_{\gamma^n, \mu}) \rightarrow S_t(\nu)$ and $C_t(\gamma^n_{\mu}) \rightarrow W_t^2(\mu, \nu)$ as $n \rightarrow \infty$ and hence

$$
(S_t(\mu) - S_t(\nu_{\gamma^n, \mu}) - \frac{K}{2} W_t^2(\mu, \nu_{\gamma^n, \mu}))^+ \frac{C_t(\gamma^n_{\mu})}{\sqrt{C_t(\gamma^n_{\mu})}} = \lim_{n \rightarrow \infty} \frac{(S_t(\mu) - S_t(\nu_{\gamma^n, \mu}) - \frac{C_t(\gamma^n_{\mu})}{2} C_t(\gamma^n_{\mu}))^+}{\sqrt{C_t(\gamma^n_{\mu})}},
$$

which shows $\leq$. \qed

We get the following as consequence of formula (52), cf. [11, Corollary 13].

Corollary 4.5. Suppose that $S$ is $K$-convex. Then for every $t \in [0,T]$

$$
|\nabla_t S_t|^2(\mu) \leq \liminf_n |\nabla_t S_t|^2(\mu_n),
$$

whenever $\mu_n \rightharpoonup \mu$ as $n \rightarrow \infty$ such that $\sup_n S_t(\mu_n) < \infty$. Further $\mu \mapsto |\nabla_t S_t|^2$ is convex with respect to linear interpolation on the sublevels of $S$.

Proof. Consider the map $\mu \mapsto C_t(\gamma_{\mu})$. It is clearly linear. Also, one can show that it is weakly continuous on sublevels of the entropy. From [11 Proposition 11] we further know that $\mu \mapsto S_t(\mu) - S_t(\nu_{\gamma, \mu})$ is lower semicontinuous with respect to weak convergence on sublevels of the entropy and convex with respect to linear interpolation. Hence

$$
\mu \mapsto S_t(\mu) - S_t(\nu_{\gamma, \mu}) - \frac{K}{2} C_t(\gamma_{\mu})
$$

is lower semicontinuous with respect to weak convergence on the sublevels of the entropy. The same holds true for its positive part. Now apply that the function
Ψ: \( \mathbb{R}^2 \rightarrow \mathbb{R} \) defined by

\[
\Psi(a, b) = \begin{cases} 
\frac{a^2}{b} & \text{if } b > 0, \\
0 & \text{if } a = b = 0, \\
+\infty & \text{if } a \neq 0, b = 0 \text{ or } b < 0, 
\end{cases}
\]

is convex, continuous on \([0, \infty)^2 \setminus \{(0, 0)\}\) and increasing in \(a\), and the conclusion follows. From formula (52) the assertion follows.

\[\blacksquare\]

### 4.3 Existence and Uniqueness of EDE-gradient flow for the entropy

In this section we want to show existence and uniqueness of the dynamic EDI-gradient flow with respect to the functional \(S\) on the complete geodesic space \((\mathcal{P}_2(X), W_t)\). For this we additionally have to assume that \(X\) is boundedly compact, i.e., closed balls are compact. For this reason we can take A5 for granted, as shown in the next lemma.

**Lemma 4.6.** Assume that \(X\) is boundedly compact. Then the following holds true. If \((\mu_n) \subset \mathcal{P}_2(X)\) with \(\sup_{n,m} W_t(\mu_n, \mu_m) < \infty\), then \((\mu_n)_n\) is sequentially precompact with respect to weak convergence.

**Proof.** If \(\sup_{n,m} W_t(\mu_n, \mu_m) < \infty\) for a sequence \((\mu_n) \subset \mathcal{P}_2(X)\) the second moments are uniformly bounded. Then Lemma 16 in [11] implies that \((\mu_n)\) is tight. Applying Prokhorov’s theorem we infer that \((\mu_n)\) is weakly sequentially precompact.

**Theorem 4.7.** Assume additionally that \(X\) is boundedly compact. Suppose that \(S\) is \(K\)-convex for some \(K \in \mathbb{R}\). Then for every \(\bar{\mu} \in \text{Dom}(S)\) there exists a curve \((\mu_t) \in C^2([0, T]; \mathcal{P}_2(X))\) starting in \(\bar{\mu}\) and satisfying

\[
S_t(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_r|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r) dr \leq S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r) dr, \quad (53)
\]

for every \(t \in [0, T]\).

**Proof.** We may apply Proposition 3.22 and obtain a limit curve \(\mu \in C^2([0, T]; \mathcal{P}_2(X))\) starting in \(\bar{\mu}\) such that

\[
\tilde{\mu}_t^{hn} \to \mu, \quad \text{and } \bar{\mu}_t^{hn} \to \mu_t \quad \forall t \in [0, T],
\]

where \(\tilde{\mu}^{hn}\) and \(\bar{\mu}^{hn}\) are defined as in Definition 3.18 and satisfy by Proposition 3.19

\[
S_{\tilde{\mu}_t^{hn}}(\tilde{\mu}_t^{hn}) + \frac{1}{2} \int_0^{\tilde{\mu}_t^{hn}(t)} (Dsp_t^{hn})^2 dr + \frac{1}{2} \int_0^{\tilde{\mu}_t^{hn}(t)} (Dst_t^{hn})^2 dr = S_0(\bar{\mu}) + \int_0^{\tilde{\mu}_t^{hn}(t)} (\partial_r S_r)(\bar{\mu}_t^{hn}) dr.
\]

31
From Corollary 4.15, Lemma 3.14 and Lemma 3.4 together with (12), applying Fatou’s Lemma we obtain

\[
\int_0^t |\nabla_r S_t|^2(\mu_r)dr \leq \liminf_{n \to \infty} \int_0^t |\nabla_r S_t|^2(\tilde{\mu}_r^{h_n})dr
\]

\[
\leq \liminf_{n \to \infty} \int_0^t \left( |\nabla_{h_n(r)} S_t| |(\tilde{\mu}_r^{h_n}) + |\nabla_r S_t| |(\tilde{\mu}_r^{h_n}) - |\nabla_{h_n(r)} S_t| |(\tilde{\mu}_r^{h_n})|^2 \right)dr
\]

\[
\leq \liminf_{n \to \infty} \left[ \int_0^t (Dsl_r^{h_n})^2 dr + 2Ch_n \int_0^t (Dsl_r^{h_n})^2 dr + Ch_n^2 \int_0^t (Dsl_r^{h_n})^2 dr \right].
\]

We deduce

\[
\int_0^t |\nabla_r S_t|^2(\mu_r)dr \leq \liminf \int_0^t (Dsl_r^{h_n})^2 dr
\]

from the estimate \(\int_0^t (Dsl_r^{h_n})^2 dr \leq S_0(\mu) + L^*T - \inf_{t,\mu} S_t(\mu)\).

To show that (53) is valid, it is left to show that

\[
\lim_{n \to \infty} \int_0^t (\partial_r S_t)(\tilde{\mu}_r^{h_n})dr = \int_0^t (\partial_r S_t)(\mu_r)dr.
\]

This already follows if we prove that a stronger convergence than weak convergence of measures holds true. In fact, from (45) we know that there exists a subsequence \(\tilde{\mu}_r^{h_n}\) (cf. [6, Theorem 4.5.9]), the Dunford-Pettis Theorem ([6, Corollary 4.7.19]) ensures that there exists a subsubsequence \(\tilde{\mu}_r^{h_nk}\) that converges in the weak topology of \(L^1(X, m)\) to the function \(\rho_t \in L^1(X, m)\). Hence for the original subsequence we already have

\[
\tilde{\mu}_r^{h_n} \to \rho_t \text{ in } L^1(X, m) \quad \forall t \in [0, T].
\]

As a direct consequence we obtain (53), since similar as in Proposition 3.23

\[
S_t(\mu_t) + \frac{1}{2} \int_0^t |\tilde{\mu}|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_t|^2(\mu_r)dr
\]

\[
\leq \liminf_{n \to \infty} \left[ S_{h_n(t)}(\tilde{\mu}^{h_n}) + \frac{1}{2} \int_0^t (Dsl_r^{h_n})^2 dr + \frac{1}{2} \int_0^t (Dsl_r^{h_n})^2 dr \right]
\]

\[
\leq \liminf_{n \to \infty} \left[ S_0(\tilde{\mu}) + \int_0^t (\partial_r S_t)(\tilde{\mu}^{h_n})dr \right]
\]

\[
\leq S_0(\tilde{\mu}) + \int_0^t (\partial_r S_t)(\mu_r)dr.
\]
Remark 4.8. Actually, the statement of Theorem 4.7 holds true without assuming that $X$ is boundedly compact, since $m$ is assumed to be finite. If $(X,d)$ is a Polish space and $m \in \mathcal{P}(X)$ we may apply $z \log z \geq -1/e$ and Jensen’s inequality to obtain

$$\text{Ent}_m(\mu) \geq \mu(E) \log \left( \frac{\mu(E)}{m(E)} \right) - \frac{1}{e} \quad \forall E \in \mathcal{B}(X).$$

Taking into account that the singleton $\{m\}$ is tight, this shows tightness of the sublevels of the entropy since $\mu(E) \to 0$ as $m(E) \to 0$. Hence we could replace our assumption A5 in section 3.3 by assuming that for each $t \in [0,T]$ the sublevels of the functional are sequentially $\sigma$-compact. See also [5, Remark 7.3].

Theorem 4.9. Assume $S$ is $K$-convex and $\bar{\mu} \in \text{Dom}(S)$. Then there exists at most one dynamic EDI-gradient flow. Moreover we have equality in (53), i.e. for every $t \in [0,T]$ it holds the following dynamic EDE

$$S_t(\mu_n) + \frac{1}{2} \int_0^t |\dot{\mu}_r|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r) dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r) dr.$$  

Proof. Let us first observe that a weak chain rule for gradient flows is applicable. For this we prove that a variant of the assumption in Proposition 3.7 concerning the time derivative is satisfied by the entropy $S_t$. We choose a sequence $\mu_n = \rho_n m$ converging to $\mu = \rho m$ such that $\sup_n S(\mu_n) < \infty$. We need to show that for almost every $t$

$$\lim_{n \to \infty} \frac{S_{t_n}(\mu_n) - S_t(\mu_n)}{t_n - t} = \lim_{n \to \infty} \int_X \frac{f_{t_n}(x) - f_t(x)}{t_n - t} \rho_n dm(x) = \partial_t S_t(\mu),$$

if $t_n \searrow t$ as $n \to \infty$. This would imply the weak chain rule (22) in Proposition 3.7 restricted to curves which are contained in the sublevels of the functional. In order to show (54) note that as in the proof of Theorem 4.7 the sequence $(\rho_n)$ is equi-integrable and thus $\rho_n$ converges to $\rho$ in duality with $L^\infty$ functions. Then we decompose

$$\int_X \frac{f_{t_n} - f_t}{t_n - t} \rho_n dm = \int_X \left( \frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t \right) \rho_n dm + \int \partial_t f_t \rho_n dm$$
$$= \int_{|\rho_n| \leq M} \left( \frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t \right) \rho_n dm + \int_{|\rho_n| \geq M} \left( \frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t \right) \rho_n dm$$
$$+ \int \partial_t f_t \rho_n dm.$$ 

The third integral clearly converges to $\int \partial_t f_t \rho dm = \partial_t S_t(\mu)$ by Lemma 4.2 while the first vanishes by dominated convergence. The second vanishes after letting $n \to \infty$ and then $M \to \infty$ by equi-integrability of $(\rho_n)$.

Let us assume that there exist two dynamic EDI-gradient flows $(\mu^1_t)$, $(\mu^2_t)$ starting from $\bar{\mu} \in \text{Dom}(S)$. As seen in the proof of Theorem 4.7 we know that these curves are contained in the sublevels of $S$ and hence together with the weak chain rule it follows

$$S_t(\bar{\mu}) = S_t(\mu^1_t) + \frac{1}{2} \int_0^t |\dot{\mu^1}_r|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu^1_r) dr - \int_0^t \partial_r S_r(\mu^1_r) dr,$$
Now define
\[ S_0(\mu) = S_t(\mu^2) + \frac{1}{2} \int_0^t |\mu_r^2|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r^2) dr - \int_0^t \partial_r S_r(\mu_r^2) dr. \]

But \( \partial - \) is a dynamic forward EVI (contradicts (22)).

Then \( \mu_0^{1,2} = \bar{\mu} \) and from the strict convexity of the entropy, the convexity of the squared slope (Corollary 4.3), the convexity of the squared speed (Lemma 4.1) and the linearity of \( \partial_r S_r \) (Lemma 4.2) we have that
\[ S_0(\mu) > S_t(\mu^{1,2}) + \frac{1}{2} \int_0^t |\mu_r^{1,2}|^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r^{1,2}) dr - \int_0^t \partial_r S_r(\mu_r^{1,2}) dr, \]
whenever these curves are different. But since (54) is applicable to \( \mu^{1,2} \), this contradicts (22).

5 Dynamic gradient flows in Hilbert spaces

Let \( H \) be a separable Hilbert space with a family of scalar products \( \langle \cdot, \cdot \rangle_t \). We assume that (12) holds for the distances \( ||x - y||_t := \sqrt{\langle x - y, x - y \rangle_t} \).

Let \( E : [0, T] \times H \to \mathbb{R} \cup \{+\infty\} \) be a functional such that \( x \mapsto E_t(x) \) is convex and lower semicontinuous. Again we require that the domain \( \text{Dom}(E_t) = \{ x : E_t(x) < \infty \} \) is time independent. The subdifferential \( D_t^- E_t(x) \) of \( E_t \) at some \( x \in \text{Dom}(E) \) is the set of all \( v \in H \) such that
\[ E_t(y) - E_t(x) \geq \langle v, y - x \rangle_t, \quad \forall y \in H. \]

It follows from the definition of the subdifferential that \( D_t^- E_t \) is monotone, i.e. for every \( v \in D_t^- E_t(x) \), \( w \in D_t^- E_t(y) \) we have
\[ \langle v - w, x - y \rangle_t \geq 0. \] (55)

Note that \( D_t^- E_t(x) \) is closed and convex. Hence we can set \( \nabla_t E_t(x) \) as the element of minimal \( ||\cdot||_t \)-norm in \( D_t^- E_t(x) \) as soon as \( D_t^- E_t(x) \neq \emptyset \).

**Definition 5.1.** We say that \( (x_t) \) is a dynamic gradient flow for \( E_t \) starting from \( x \in H \) if it is locally absolutely continuous and
\[ \partial_t x_t \in -D_t^- E_t(x_t) \quad \text{for a.e. } t > 0 \]
and \( \lim_{t \downarrow 0} x_t = x \).

We cannot hope to have a minimal selection result, i.e. \( \frac{d^+}{dt} x_t = -\nabla_t E_t(x_t) \).

We illustrate this in the following example.

**Example 6.** Consider once again the energy functional \( E_t(x) = |x - t| \) on \( \mathbb{R} \). Then the curve \( x_t = t \) defines a gradient flow of \( E_t \) since \( \partial_t x_t = 1 \in -D_t^- E_t(x_t) \), but \( \partial_t x_t \neq -\nabla E_t(x_t) = 0 \).

In the following we show that the gradient flow in the sense of Definition 5.1 is a dynamic forward EVI \((-L/2, \infty) \) gradient flow introduced in Section 8.2.

We recall that for \( s, t \in [0, T], \gamma \in AC^2([0, 1]; H) \) the action of the curve
\[ A_{s, t}(\gamma) := \lim_{h \to 0} \sup \left\{ \sum_{i=1}^n (a_i - a_{i-1})^{-1} ||\gamma_{a_{i-1}} - \gamma_{a_i}||^2_{\gamma + a_i(t-s)} \right\}, \]
where the supremum runs over all partitions $0 = a_0 < a_1 < \cdots a_n = 1$ such that $a_i - a_{i-1} \leq h$ for some $h > 0$.

For $x, y \in H$ we define

$$||x - y||_{s,t}^2 := \inf A_{s,t}(\gamma),$$

where the infimum runs over all curves $\gamma \in AC^2([0,1];H)$ such that $\gamma_0 = x$ and $\gamma_1 = y$.

**Proposition 5.2.** Let $E : [0,T] \times H \to (-\infty, +\infty]$ be a functional such that $x \mapsto E_t(x)$ is convex and lower semicontinuous for each $t \in [0,T]$. Let $(x_t)$ be a gradient flow of $E$ in the sense of Definition 5.1. Then, with $L$ denoting the the logarithmic Lipschitz control of the distances, $(x_t)$ is a dynamic forward EVI$(L/2, \infty)$ gradient flow, i.e. for all $y \in \text{Dom}(E)$ and a.e. $t$

$$\frac{1}{2} \partial_t^+ ||x_t - y||_{s,t}^2 - \frac{L}{4} ||x_t - y||_t^2 \leq E_t(y) - E_t(x_t).$$

**Proof.** Let $y \in \text{Dom}(E)$. Then

$$\frac{1}{2} \partial_t^+ ||x_t - y||_{s,t}^2 \leq \limsup_{s \to t} \frac{1}{2(s-t)} \left( ||x_s - y||_{s,t}^2 - ||x_t - y||_{s,t}^2 \right)$$

$$\leq \limsup_{s \to t} \frac{1}{2(s-t)} \left( ||x_s - y||_{s,t}^2 - ||x_t - y||_{s,t}^2 \right)$$

$$+ \limsup_{s \to t} \frac{1}{2(s-t)} \left( ||x_t - y||_{s,t}^2 - ||x_t - y||_{s,t}^2 \right)$$

The first limsup can be estimated with the help of Proposition 7.2(iii) in [15]

$$\limsup_{s \to t} \frac{1}{2(s-t)} \left( ||x_s - y||_{s,t}^2 - ||x_t - y||_{s,t}^2 \right)$$

$$\leq \limsup_{s \to t} \frac{1}{2(s-t)} \left( \frac{e^{L(t-s)} - 1}{L|t-s|} \right) ||x_s - y||_{s,t}^2$$

$$= \limsup_{s \to t} \frac{1}{2(s-t)} \left( \frac{1}{2} \frac{L|t-s|^2 + o(|t-s|^2)}{L|t-s|} \right) ||x_s - y||_{s,t}^2$$

$$\leq \frac{L}{4} ||x_t - y||_t^2,$$

where the last inequality follows from the continuity of $t \mapsto x_t$ and $t \mapsto ||x_t||$.

For the second limsup we apply that $(x_t)$ is supposed to be a gradient flow of $E$:

$$\limsup_{s \to t} \frac{1}{2(s-t)} \left( ||x_s - y||_{s,t}^2 - ||x_t - y||_{s,t}^2 \right)$$

$$= \langle x_t - y, \partial_t x_t \rangle \leq E_t(y) - E_t(x_t)$$

for a.e. $t \geq 0$. Combining these two observations we conclude

$$\frac{1}{2} \partial_t^+ ||x_t - y||_{s,t}^2 \leq \frac{L}{4} ||x_t - y||_t^2 + E_t(y) - E_t(x_t),$$

which proves the claim. \( \square \)
5.1 Existence and Uniqueness

We assume that the following holds for the energy functional, cf. [22].

1. \( x \mapsto E_t(x) \) is lower semicontinuous and \( E_t(x) \geq 0 \) \( \forall (t, x) \in [0, T] \times \text{Dom}(E) \).

2. \( \exists C_1 \forall x \in \text{Dom}(E) \forall s, t \in [0, T] : |E_t(x) - E_s(x)| \leq C_1 |t - s| \).

By virtue of the functional’s lower semicontinuity we obtain that if \( v_n \in D^- E_t(x_n) \) and \( x_n \to x \), \( v_n \rightharpoonup v \), then \( v \in D^- E_t(x) \).

We write \( e(x) := \sup_{t \in [0, T]} E_t(x) \).

Note that from the Lipschitz property it follows that there exists a constant \( C_2 > 0 \) such that for all \( x \in \text{Dom}(E) \)

\[
e(x) \leq C_2 \inf_{t \in [0, T]} E_t(x).
\]

(56)

Approximation

We fix a time step \( h > 0 \) and subdivide the interval \([0, T]\) into

\[ \mathcal{P}_h := \{ t_0 = 0 < t_1 < \cdots < t_{N-1} < T \leq t_N \}, \quad t_n = nh, N \in \mathbb{N}. \]

For \( 0 \leq t \leq T \) we define the piecewise constant interpolants \( \tilde{h}(t) \) and \( h(t) \) associated with the partition \( \mathcal{P}_h \) in the following way;

\[ \tilde{h}(0) = h(0), \text{ and for } t \in (t_{n-1}, t_n] \quad \tilde{h}(t) = t_n, \quad h(t) = t_{n-1}. \]

(57)

The definition implies that \( \tilde{h}(t) \searrow t \) and \( h(t) \not\nearrow t \) if \( h \searrow 0 \).

For a given initial value \( \tilde{x} \) we recursively define a sequence \( (x^h_n) \) of minimizers by

\[
x^h_0 := \tilde{x}, \quad x^h_n := \arg \min_x \left\{ E_{t_n}(x) + \frac{1}{2h} ||x - x^h_{n-1}||^2_{t_n} \right\}.
\]

(58)

We can argue as in the proof of Proposition 3.12 and directly obtain for every \( \tilde{x} \in \text{Dom}(E) \) and \( h > 0 \) a (unique) solution to the minimization problem (55). As in Section 5.3 we define piecewise constant interpolants by setting

\[ \tilde{x}^h_t := x^h_n \text{ for } t \in (t_{n-1}, t_n], \quad \tilde{x}^h_t := x^h_{n-1} \text{ for } t \in (t_{n-1}, t_n), \]

and moreover, the piecewise linear interpolant

\[ \tilde{x}^h_t = \frac{t - t_{n-1}}{h} x^h_n + \frac{t_n - t}{h} x^h_{n-1} \text{ for } t \in [t_{n-1}, t_n). \]

For \( t \in (t_{n-1}, t_n) \) we denote the time derivative of \( t \mapsto \tilde{x}^h_t \) by \( \dot{\tilde{x}}^h_t \).

Recall that the variational interpolation is a map \( t \mapsto \tilde{x}^h_t \) defined by

\[ \dot{\tilde{x}}^h_t = \arg \min_x \left\{ E_t(x) + \frac{1}{2r} ||x - x^h_{n-1}||^2_{t_n} \right\}, \]

for \( t = t_{n-1} + r \in (t_{n-1}, t_n] \),

36


Indeed, arguing similar as in (39) we obtain that for $r_h$.

$$\tilde{v}^h_t := \frac{x^h_t - x^h_{n-1}}{t - t_{n-1}} \quad \forall t \in (t_{n-1}, t_n].$$

As in Section 5.3, in order to extract a converging subsequence, we prove a priori estimates on the discrete solutions. The proof is along the lines of Proposition 6.3 in [21].

**Proposition 5.3.** The following inequality holds for the interpolants $\tilde{x}^h$, $x^h$, $\bar{x}^h$ and $\bar{v}^h$

$$E_{\mathcal{H}(t)}(x^h_t) + \frac{1}{2} \int_{\mathcal{H}(s)}^{\mathcal{H}(t)} ||\tilde{x}^h_r||^2_{H(r)} dr + \frac{1}{2} \int_{\mathcal{H}(s)}^{\mathcal{H}(t)} ||\bar{v}^h_r||^2_{H(r)} dr \leq E_{\mathcal{H}(s)}(\tilde{x}^h_s) + C_1 \int_{\mathcal{H}(s)}^{\mathcal{H}(t)} e(\bar{v}^h_r) dr.$$  \hspace{1cm} (59)

In particular there exists a constant $M$ such that for all $h > 0$

$$\sup_{t \in (0, T)} e(\tilde{x}^h_t) \leq M, \quad \sum_{n=0}^{N} \frac{1}{h} ||x^h_n - x^h_{n-1}||^2_{H_n} \leq M,$$  \hspace{1cm} (60)

$$\int_0^T ||\tilde{x}^h_r||^2_{H(r)} dr \leq M, \quad \int_0^T ||\bar{v}^h_r||^2_{H(r)} dr \leq M.$$  \hspace{1cm} (61)

Moreover

$$||\tilde{x}^h_t - x^h_t||^2 \in O(h), \quad ||x^h_t - \bar{x}^h_t||^2 \in O(h), \quad ||\tilde{x}^h_t - \bar{x}^h_t||^2 \in O(h).$$  \hspace{1cm} (62)

**Proof.** Consider the map

$$r \mapsto J_s, r(y) := \min_x \left\{ E_{s+r}(x) + \frac{1}{2r} ||x - y||^2_{s+h} \right\}$$

for a given $s \in [0, T]$, $y \in D$, $0 < r < T - s$. We claim that this map is differentiable almost everywhere in $(0, T - s)$ and for every $r_0 \in (0, T - s)$ for the minimizer $(0, r_0] \ni r \mapsto x_r$ it holds

$$\frac{1}{2r_0} ||x_{r_0} - y||^2_{s+h} + \int_0^{r_0} \frac{1}{2r^2} ||x_r - y||^2_{s+h} dr + E_{s+r_0}(x_{r_0}) \leq E_s(y) + C_1 \int_0^{r_0} e(x_r) dr.$$  \hspace{1cm} (63)

Indeed, arguing similar as in [21] we obtain that for $r_1 < r_2 \in (0, T - s)$

$$J_{s,r_2}(y) - J_{s,r_1}(y) - (E_{s+r_2}(x_{r_1}) - E_{s+r_1}(x_{r_1})) \leq - \frac{1}{2r_1 r_2} (r_2 - r_1) ||x_{r_1} - y||^2_{s+h} \leq 0,$$  \hspace{1cm} (64)

hence the map $r \mapsto J_s, r(y)$ is the sum of a locally Lipschitz and of a nonincreasing function

$$J_{s,r_2}(y) \leq J_{s,r_1}(y) + (r_2 - r_1) C_1 e(x_{r_1}),$$

37
and differentiable almost everywhere. So let \( r \in (0, T - s) \) be a differentiable point of \( r \mapsto J_{s,r}(y) \). Then with (64) we get

\[
\frac{d}{dr} J_{s,r}(y) + \frac{1}{2r^2} \| x_r - y \|_{s+h}^2 \\
= \lim_{h \to 0} \left( \frac{J_{s,r+h}(y) - J_{s,r}(y)}{h} + \frac{1}{2(r+h)^2} \| x_r - y \|_{s+h}^2 \right) \\
\leq \liminf_{h \to 0} \frac{E_{s+r+h}(x_r) - E_{s+r}(x_r)}{h} \leq C_1 e(x_r),
\]

and integrating from 0 to \( r_0 \) gives us (63).

Applying (63) with \( t \in (t_{n-1}, t_n) \), \( y = x_{n-1}^h \), \( s = t_{n-1} \) and \( r_0 = t - s \) we obtain for \( x_t^h \)

\[
\frac{1}{2(t - t_{n-1})} \| \tilde{x}_t^h - x_{n-1}^h \|_{t_n}^2 + \int_{t_{n-1}}^t \frac{1}{2(r - t_{n-1})^2} \| \tilde{x}_r^h - x_{n-1}^h \|_{t_n}^2 dr + E_t(\tilde{x}_r^h) \\
\leq E_{t_{n-1}}(x_{n-1}^h) + C_1 \int_{t_{n-1}}^t e(\tilde{x}_r^h) dr.
\]

Inserting \( t = t_n \) we get for the interpolants \( x_n^h \), \( \tilde{v}_n^h \)

\[
\frac{1}{2} \int_{t_{n-1}}^{t_n} \| \tilde{x}_r^h \|_{t_n}^2 dr + \int_{t_{n-1}}^{t_n} \frac{1}{2} \| \tilde{v}_r^n \|_{t_n}^2 dr + E_{t_n}(\tilde{x}_{t_n}^h) \\
\leq E_{t_{n-1}}(x_{n-1}^h) + C_1 \int_{t_{n-1}}^{t_n} e(\tilde{x}_r^h) dr.
\]

Summing over the partition we end up with (65).

Note that the minimality and (60) imply that for \( r \in (t_{n-1}, t_n) \), \( t = t_{n-1} + r \)

\[
e(\tilde{x}_{t_{n-1}}) \geq E_t(\tilde{x}_{t_{n-1}}) \geq \frac{1}{2h} \| \tilde{x}_r^h - \tilde{x}_{t_{n-1}} \|_{t_n}^2 + E_t(\tilde{x}_r^h) \geq E_t(\tilde{x}_r^h) \geq \frac{1}{C_2} e(\tilde{x}_r^h),
\]

and hence with (65) we can estimate

\[
E_{t_n}(\tilde{x}_{t_n}^h) \leq E_{t_{n-1}}(x_{n-1}^h) + C_1 C_2 \int_{t_{n-1}}^{t_n} e(\tilde{x}_{t_{n-1}}) dr.
\]

Summing over the partitions and applying (60) once more we obtain for some constant \( C > 0 \)

\[
e(\tilde{x}_{t_n}^h) \leq C(E_0(\tilde{x}_0^h) + \int_0^{t_n} e(\tilde{x}_r^h) dr).
\]

We obtain the first inequality in (60) by applying a discrete Gronwall argument (see e.g., [22, Lemma 4.5]). It directly follows that the right-hand side of (65) is bounded and (61) holds. The second inequality in (60) is a direct consequence of the first estimate in (61).

In order to show the first statement in (62) recall that (65) together with (60) implies (with some different constant \( M \))

\[
\| \tilde{x}_t^h - \tilde{x}_r^h \| \leq 2hM.
\]
The other two assertions in (62) follow from (61) via Hölder’s inequality

\[ ||x_t^h - x_s^h|| \leq \int_s^t ||\dot{x}_r^h|| dr \leq \sqrt{M(t-s)} \quad \forall 0 < s < t < T. \]

The following result provides the compactness of the approximate solutions.

**Proposition 5.4.** For every sequence of time-steps \((h_j)_{j \in \mathbb{N}}\) such that \(h_j \to 0\) as \(j \to \infty\) there exists a subsequence \(h_j\) (not relabeled) and an absolutely continuous curve \((x_t) \subset AC^2([0,T];H)\) and such that

\[ x_t^{h_j} \to x_t \text{ in } C^0([0,T];H), \]

and

\[ x_t^{h_j} \to \dot{x}_t \text{ in } L^2([0,T];H). \]

Moreover for each \(t \), \(x_t^{h_j} \to x_t \) in \(H\).

**Proof.** Let \(0 < g, h \ll T\) be two stepsizes and \({t^n}^N_{n=0}\), \({t^n}^{N_h}_{n=0}\) the corresponding partitions of the interval \([0,1]\). Let \({x^n}^{N_h}_{n=0}\) and \({x^n}^{N_g}_{n=0}\) be the solution to the minimizing problem (58) with respect to the stepsizes \(h\) and \(g\) respectively with initial condition \(x_0^h\) and \(x_0^g\). The Euler-Lagrange equation of \(x^h_n\) is

\[ \frac{x^h_n - x^h_{n-1}}{h} \in -D^-_{t^h_n}E_{t^h_n}(x^h_n), \]

i.e.

\[ h^{-1}(x^h_n - x^h_{n-1}, x^h_n - y)_{t^h_n} + E_{t^h_n}(x^h_n) - E_{t^h_n}(y) \leq 0 \quad \forall y \in H. \]

Inserting the definition of the piecewise linear interpolation \(x_t^h\) at \(t \in (t^h_{n-1}, t^h_n)\)

\[ \langle \dot{x}_t^h, x^h_t - y \rangle_{t^h_n} + E_{t^h_n}(x^h_t) - E_{t^h_n}(y) \leq (t - t^h_n) \left( ||\dot{x}^h_t||^2_{t^h_n} + \frac{1}{h}(E_{t^h_n}(x^h_n) - E_{t^h_n}(x^h_{n-1})) \right) \quad \forall y \in H, \quad (67) \]

where we applied the convexity of \(x \mapsto E_t(x)\). The same argumentation for \(x_t^g\) at \(t \in (t^g_{m-1}, t^g_m)\) yields

\[ \langle \dot{x}_t^g, x^g_t - y \rangle_{t^g_m} + E_{t^g_m}(x^g_t) - E_{t^g_m}(y) \leq (t - t^g_m) \left( ||\dot{x}^g_t||^2_{t^g_m} + \frac{1}{g}(E_{t^g_m}(x^g_m) - E_{t^g_m}(x^g_{m-1})) \right) \quad \forall y \in H. \quad (68) \]

For \(t \in (t^h_{n-1}, t^h_n) \cap (t^g_{m-1}, t^g_m)\) we get by putting \(y = x_t^g\) into (67) and \(y = x_t^h\) into (68) and adding them

\[ \langle \dot{x}_t^h, x^h_t - x^g_t \rangle_{t^h_n} + \langle \dot{x}_t^g, x^g_t - x^h_t \rangle_{t^g_m} + E_{t^h_n}(x^h_t) - E_{t^g_m}(x^g_t) + E_{t^h_n}(x^h_t) - E_{t^h_n}(x^g_{m-1}) \leq (t - t^h_n) \left( ||\dot{x}^h_t||^2_{t^h_n} + \frac{E_{t^h_n}(x^h_n) - E_{t^h_n}(x^h_{n-1})}{h} \right) \]

\[ + (t - t^g_m) \left( ||\dot{x}^g_t||^2_{t^g_m} + \frac{E_{t^g_m}(x^g_m) - E_{t^g_m}(x^g_{m-1})}{g} \right). \quad (69) \]
The Lipschitz property (12) of the metric together with the polarization identity gives
\[
\langle \dot{x}^h_t, x^h_t - x^q_t \rangle_{\mathbb{E}^h_t} + \langle \dot{x}^q_t, x^h_t - x^q_t \rangle_{\mathbb{E}^q_t}
\geq \langle \dot{x}^h_t - \dot{x}^q_t, x^h_t - x^q_t \rangle_{\mathbb{E}^h_t} - L(t^h_m - t^q_m)\left(\langle \dot{x}^h_t, x^h_t - x^q_t \rangle_{\mathbb{E}^h_t} + \frac{1}{2}||\dot{x}^h_t - (x^h_t - x^q_t)||^2_{\mathbb{E}^h_t}\right),
\]
while the Lipschitz property of the energy yields
\[
E_{t^h}(x^h_t) - E_{t^q}(x^q_t) \geq -C_1|t^h_m - t^q_m|\left(E_{t^h}(x^h_t) + E_{t^q}(x^q_t)\right).
\]
Inserting these two inequalities into (69) we find
\[
\frac{d}{dt}||x^h_t - x^q_t||^2_{\mathbb{E}^h_t} = \langle \dot{x}^h_t - \dot{x}^q_t, x^h_t - x^q_t \rangle_{\mathbb{E}^h_t}
\leq (t - t^h_m)\left(||\dot{x}^h_t||^2_{\mathbb{E}^h_t} + E_{t^h}(x^h_{t^h_m}) - E_{t^h}(x^h_{t^h_m-1})\right)
+ (t - t^q_m)\left(||\dot{x}^q_t||^2_{\mathbb{E}^q_t} + E_{t^q}(x^q_{t^q_m}) - E_{t^q}(x^q_{t^q_m-1})\right)
+ L|t^h_m - t^q_m|\left(\langle \dot{x}^h_t, x^h_t - x^q_t \rangle_{\mathbb{E}^h_t} + \frac{1}{2}||\dot{x}^h_t - (x^h_t - x^q_t)||^2_{\mathbb{E}^h_t}\right)
+ C_1|t^h_m - t^q_m|\left(E_{t^h}(x^h_t) + E_{t^q}(x^q_t)\right).
\]
Integrating \((70)\) on the interval \((t^h_{m-1} \cup t^q_{m-1}, t)\) we can estimate
\[
||x^h_t - x^q_t||^2_{\mathbb{E}^h_t} - ||x^h_{t^h_{m-1} \cup t^q_{m-1}} - x^q_{t^h_{m-1} \cup t^q_{m-1}}||^2_{\mathbb{E}^h_t}
\leq h \int_{t^h_{m-1} \cup t^q_{m-1}}^t \left(-||\dot{x}^h_t||^2_{\mathbb{E}^h_t} + \frac{E_{t^h}(x^h_{t^h_m}) - E_{t^h}(x^h_{t^h_m-1})}{h}\right)dr
+ g \int_{t^h_{m-1} \cup t^q_{m-1}}^t \left(-||\dot{x}^q_t||^2_{\mathbb{E}^q_t} + \frac{E_{t^q}(x^q_{t^q_m}) - E_{t^q}(x^q_{t^q_m-1})}{g}\right)dr
+ L(h \land g) \int_{t^h_{m-1} \cup t^q_{m-1}}^t \left(||\dot{x}^h_t||^2_{\mathbb{E}^h_t} + ||\dot{x}^q_t||^2_{\mathbb{E}^q_t}\right)dr
+ C_1(h \land g) \int_{t^h_{m-1} \cup t^q_{m-1}}^t \left(E_{t^h}(x^h_t) + E_{t^q}(x^q_t)\right)dr.
\]
Summing over the partition \(\{t^h_{j=0}^{N_h+1}\}_j \cup \{t^q_{m=0}^{N_q}\}_m\) and exploiting
the Lipschitz property of $t \mapsto || \cdot ||_t$

$$
||x^h_t - x^q_t||^2_{\ell^n_{t}} \leq ||x^h_0 - x^q_0||^2_{\ell^n_{0}} + L(h \wedge g) \sum_{j=1}^{n+m} ||x^h_{t^n_{j-1}} - x^q_{t^n_{j-1}}||^2_{g(t^n_{j-1})}
$$

$$
+ \sum_{j=1}^{n} \int_{t^n_{j-1}}^{t^n_{j}} (-h||\dot{x}^h_r||^2_{p^n_{j}} + E_{t^n_{j}}(x^h_{t^n_{j-1}}) - E_{t^n_{j}}(x^q_{t^n_{j-1}})) \, dr
$$

$$
+ \sum_{j=1}^{m} \int_{t^n_{j-1}}^{t^n_{j}} (-g||\dot{x}^q_r||^2_{p^n_{j}} + E_{t^n_{j}}(x^q_{t^n_{j-1}}) - E_{t^n_{j}}(x^q_{t^n_{j-1}})) \, dr
$$

$$
+ L(h \wedge g) \int_0^t (||\dot{x}^h_r||^2_{g(t(r))} + ||\dot{x}^q_r||^2_{g(t(r))}) \, dr
$$

$$
+ C_1(h \wedge g) \int_0^t (E_{h(t(r))}(x^h_r) + E_{h(t(r))}(x^q_r)) \, dr.
$$

Applying once more the Lipschitz property of $t \mapsto E_t(x)$, we can further estimate

$$
||x^h_t - x^q_t||^2_{\ell^n_{t}} \leq ||x^h_0 - x^q_0||^2_{\ell^n_{0}} + L(h \wedge g) \sum_{j=1}^{n+m} ||x^h_{t^n_{j-1}} - x^q_{t^n_{j-1}}||^2_{g(t^n_{j-1})}
$$

$$
+ h(E_0(x^h_0) - E_n(x^h_n)) + g(E_0(x^q_0) - E_m(x^q_m))
$$

$$
+ L(h \wedge g) \int_0^t ||\dot{x}^h_r||^2_{h(t(r))} \, dr + C(h \wedge g) \int_0^t ||\dot{x}^h_r - \dot{x}^q_r||^2_{g(t(r))} \, dr
$$

$$
+ C_1(h \wedge g) \int_0^t (E_{h(t(r))}(x^h_r) + E_{h(t(r))}(x^q_r)) \, dr
$$

$$
+ C_1 \sum_{j} \int_{t^n_{j-1}}^{t^n_{j}} (t^n_{j} - t^n_{j-1})E_{t^n_{j}}(x^q_{t^n_{j-1}}) \, dr + C_1 \sum_{k} \int_{t^n_{k-1}}^{t^n_{k}} (t^n_{k} - t^n_{k-1})E_{t^n_{k}}(x^q_{t^n_{k-1}}) \, dr.
$$

From the positivity of $E$ and from (61) as well as (60) we can deduce (with varying constants)

$$
||x^h_t - x^q_t||^2_{\ell^n_{t}} \leq ||x^h_0 - x^q_0||^2_{\ell^n_{0}} + L(h \wedge g) \sum_{j=1}^{n+m} ||x^h_{t^n_{j-1}} - x^q_{t^n_{j-1}}||^2_{g(t^n_{j-1})}
$$

$$
+ hE_0(x^h_0) + gE_0(x^q_0)
$$

$$
+ L(h \wedge g) \int_0^t ||\dot{x}^h_r - \dot{x}^q_r||^2_{g(t(r))} \, dr
$$

$$
+ C(h \wedge g) + C(h \wedge g)
$$

$$
\leq ||x^h_0 - x^q_0||^2_{\ell^n_{0}} + L(h \wedge g) \sum_{j=1}^{n+m} ||x^h_{t^n_{j-1}} - x^q_{t^n_{j-1}}||^2_{g(t^n_{j-1})}
$$

$$
+ C(h \wedge g) + C(h \wedge g).
$$

The last inequality follows from

$$
\sup_{t} ||x^h_t||_{t} = \sup_{n} ||x^h_n||_{t} \leq \sup_{n}(\sqrt{2CnhM} + ||x^h_0||_{t}) \leq \sqrt{2CTM} + ||x^h_0||_{t},
$$

41
where we used the definition of \((x^h_t)_{t \in [0,T]}\) in the second equality, triangle inequality and Cauchy-Schwartz inequality in the second, and \(C = C(L)\) is the constant arising from the log-Lipschitz control \((12)\) of the metric.

For \(h, g\) sufficiently small there exists a \(\kappa\) satisfying \(1 - L(h \wedge g) \geq \frac{1}{\kappa} > 0\). Applying the discrete Gronwall lemma [22, Lemma 4.5] we finally obtain

\[
\|x^h_t - x^h_s\|^2_{L^m} \leq \kappa C(h + g)e^{\kappa(n+m)(h \wedge g)} \leq \kappa C(h + g)e^{2T\kappa}. \tag{74}
\]

This shows that if \(h_j\) is a vanishing sequence of stepsizes, \(\{x^h_{t_j}\}_{j} \subset C^0([0,T]; H)\) is a Cauchy sequence. Since \(C^0([0,T]; H)\) is a Banach space there exists a continuous curve \((x_t)_{t \in [0,T]}\) and a subsequence (not relabeled) such that \(x^h_{t_j} \to x_t\) in \(C^0([0,T]; H)\) as \(j \to \infty\). From (62) it follows immediately that also \(x^h_{t_i}, x^h_{t_j}\) converge to \(x_t\).

Since \(\int_0^T \|\dot{x}^h_t\|^2 dr \leq M\) we can extract a further subsequence (not relabeled) with \(\dot{x}^h_{t_i} \to u\) in \(L^2([0,T]; H)\) where \(u\) is some function in \(L^2([0,T]; H)\). As a consequence we obtain that the limit function \(x \in AC^2([0,T]; H)\) since for all \(0 < s < t < T\)

\[
\|x_t - x_s\|_{t^*} = \lim_{j \to \infty} \|x^h_{t_j} - x^h_{s_j}\|_{t^*} = \lim_{j \to \infty} \|\int_s^t \dot{x}^h_{t_j} dr\|_{t^*} \leq \int_s^t \|u_r\|_{t^*} dr, \tag{75}
\]

where \(t^*\) is an arbitrarily fixed timepoint in \([0,T]\). We still have to show that \(u_r = \dot{x}_r\) almost everywhere. This follows again straightforward from the weak convergence of \(\dot{x}^h_{t_i}\). Let \(y \in H\), then

\[
\langle x_t - x_s, y \rangle_{t^*} = \lim\langle x^h_{t_j} - x^h_{s_j}, y \rangle_{t^*} = \lim\langle \int_s^t \dot{x}^h_{t_j} dr, y \rangle_{t^*} = \langle \int_s^t u_r dr, y \rangle_{t^*}.
\]

Since \(y \in H\) is arbitrary we obtain

\[
x_t - x_s = \int_s^t u_r dr,
\]

and hence \(\lim_{s \to t} \frac{x_t - x_s}{t - s} = u_t\) at every Lebesgue point of \(u\).

**Theorem 5.5.** Let \(E\) be as in the beginning of this section. Then for every \(x \in Dom(E)\) there exists a unique \((x_t) \in AC^2([0,T]; H)\) with \(\lim_{t \to 0} x_t = x\) such that

\[
\partial_t x_t \in -D^-_t E_t(x_t) \text{ for a.e. } t \in (0,T).
\]

**Proof.** Recall that the minimizers of \((55)\) with \(x^0_0 := x\) satisfy the Euler-Lagrange equation, that is in terms of the subdifferential of \(E\), the piecewise linear interpolant \(\dot{x}_t^h\) and the piecewise constant interpolant \(\ddot{x}_t^h\)

\[
\langle \dot{x}_t^h, x_t^h - y \rangle_{t^*} + E_t^h(\dot{x}_t^h) - E_t^h(y) \leq 0 \quad \forall y \in H, \text{ for every } t \in (t^h_{n-1}, t^h_n).
\]

The log Lipschitz property together with the polarization identity gives then for all \(y \in H\) and almost every \(t \in [0,T]\)

\[
\langle \dot{x}_t^h, x_t^h - y \rangle_t + E_t(\dot{x}_t^h) - E_t(y) \leq Lh(|\dot{x}_t^h|_t^2 + |x_t^h - y|_t^2) + C_1 h(e(\dot{x}_t^h) + e(y)).
\]

42
Integrating this inequality over the interval \((s,t)\) for some \(0 < s < t < T\) we deduce
\[
\int_s^t \langle \dot{x}_r^h, x_r^h - y \rangle_r dr + \int_s^t E_r(\dot{x}_r^h) - E_r(y) \, dr
\leq Lh \int_s^t ||\dot{x}_r^h||^2 + ||\dot{x}_r^h - y||^2 \, dr + C_1 h \int_s^t (e(\dot{x}_r^h) + e(y)) \, dr.
\tag{76}
\]
Applying Proposition 5.4 we get existence of a subsequence and a curve \((x_t) \in AC^2([0,T];H)\) such that \(\hat{x}_t^h \rightarrow x_t\) in \(C^0([0,T];H)\) and \(\dot{x}_t^h \rightarrow \dot{x}_t\) weakly in \(L^2([0,T];H)\). Hence we get for all \(y \in \text{Dom}(E)\)
\[
\int_s^t \langle \dot{x}_r, x_r - y \rangle_r dr + \int_s^t E_r(x_r) - E_r(y) dr
\leq \liminf_{h \rightarrow 0} \int_s^t \langle \dot{x}_r^h, x_r - y \rangle_r dr + \liminf_{h \rightarrow 0} \int_s^t E_r(\dot{x}_r^h) - E_r(y) \, dr
\leq \liminf_{h \rightarrow 0} \left\{ \int_s^t \langle \dot{x}_r^h, x_r^h - y \rangle_r dr + \int_s^t E_r(\dot{x}_r^h) - E_r(y) \, dr \right\}
\leq \liminf_{h \rightarrow 0} \left\{ Lh \int_s^t ||\dot{x}_r^h||^2 + ||\dot{x}_r^h - y||^2 \, dr + C_1 h \int_s^t (e(\dot{x}_r^h) + e(y)) \, dr \right\}
\leq 0,
\]
where we applied Fatou’s Lemma and the lower semicontinuity of \(x \mapsto E_t(x)\) in the first inequality, estimate \((60)\) in the third inequality and the non-negativity of \(E_t(y)\), \((59)\) and \((61)\) in the last. Dividing by \(t-s\) and letting \(s \rightarrow t\) we infer from the Lebesgue differentiation theorem that
\[
\langle \dot{x}_t, x_t - y \rangle_t + E_t(x_t) - E_t(y) \leq 0
\]
for almost every \(t \in (0,T)\) and \(y \in X\).
Since \(\hat{x}_t^h\) converges to \(x_t\) for every \(t\) we clearly have that \(\lim_{t \rightarrow 0} x_t = x\).

Suppose there exists two absolutely continuous curves \((x_t), (\hat{x}_t)_{t \in [0,T]}\) such that for every \(y \in X\) and almost every \(t \in [0,T]\)
\[
\langle \dot{x}_t, x_t - y \rangle_t + E_t(x_t) - E_t(y) \leq 0,
\]
\[
\langle \dot{x}_t, \hat{x}_t - y \rangle_t + E_t(\hat{x}_t) - E_t(y) \leq 0
\]
with \(\lim_{t \rightarrow 0} x_t = \lim_{t \rightarrow 0} \hat{x}_t = x_0\). Inserting \(\hat{x}_t\) for \(y\) into the first inequality and \(x_t\) for \(y\) into the second we obtain by adding and using \(55\)
\[
\partial_t \frac{1}{2} ||x_s - \hat{x}_s||_t^2 |_{s=t} = (\dot{x}_t - \hat{x}_t, x_t - \hat{x}_t) \leq 0.
\]
From the log-Lipschitz continuity of the metric we deduce
\[
\partial_t \frac{1}{2} ||x_s - \hat{x}_s||_t^2 |_{s=t} \leq L ||x_t - \hat{x}_t||_t^2.
\]
Applying Gronwall’s inequality we conclude \(||x_t - \hat{x}_t||_t^2 \leq e^{2Lt} ||x_0 - \hat{x}_0||_t^2 = 0\) for almost every \(t \in [0,T]\) and hence for every \(t \in [0,T]\) by continuity. This proves uniqueness. \(\blacksquare\)
Let \((X,d_t,m_t)_{t \in [0,T]}\) be a family of Polish metric measure space. We always assume that (12) holds and that there exists a reference measure \(m \in \mathcal{P}(X)\) such that \(m_t = e^{-f_t}m\) with Borel functions \(f_t\) satisfying \(|f_t(x)| \leq C\) and
\[
|f_t(x) - f_s(x)| \leq L^*|t - s|, \quad |f_t(x) - f_t(y)| \leq Cd_t(x,y).
\]
(77)

Let us denote Cheeger’s energy by \(Ch_t : L^2(X,m_t) \to [0,\infty]\)
\[
Ch_t(u) = \frac{1}{2} \inf \left\{ \liminf_{n \to \infty} \int_X (\text{lip}_t u_n)^2 dm_t | u_n \in \text{Lip}(X), \int_X |u_n - u|^2 dm_t \to 0 \right\},
\]
where \(\text{lip}_t u\) denotes the local slope defined by
\[
\text{lip}_t u(x) := \limsup_{y \to x} \frac{|u(x) - u(y)|}{d_t(x,y)}.
\]
By making use of the minimal relaxed gradient \(\nabla_t u\) (Definition 4.2), this functional admits the integral representation
\[
Ch_t(u) = \frac{1}{2} \int_X |\nabla_t u|^2 dm_t,
\]
set equal to \(+\infty\) if \(u\) has no relaxed slope. This defines a convex and lower semicontinuous functional in \(L^2(X,m_t)\) [3, Theorem 4.5].

Lemma 6.1. Fix \(t \in [0,T]\) and let \(u \in \text{Dom}(Ch_t)\). Then, assuming (12)
\[
|\nabla_t u|_* \leq e^{L^*|t-s|}|\nabla_s u|_* \quad \text{m.a.e. in } X, \quad \forall s \in [0,T].
\]
Proof. Since \(u \in \text{Dom}(Ch_t)\) we know \(u \in \text{Dom}(Ch_s)\) as well and there exist bounded Borel Lipschitz functions \(u_n \in L^2(X,m_s)\) such that
\[
u_n \to u, \quad \text{lip}_s u_n \to |\nabla_s u|_* \quad \text{strongly in } L^2(X,m_s),
\]
see e.g. [3] Lemma 4.3 (c)]. This implies that \(e^{L^*|t-s|}|\nabla_s u|_*\) is a relaxed \(d_t\)-gradient since
\[
u_n \to u, \quad e^{L^*|t-s|}\text{lip}_s u_n \to e^{L^*|t-s|}|\nabla_s u|_* \quad \text{strongly in } L^2(X,m_t)
\]
and
\[
|\nabla_t u_n|_* \leq e^{L^*|t-s|}\text{lip}_s u_n,
\]
c.f. [3] Lemma 4.3. (a)]. Thus Lemma 4.4 in [3] yields the assertion.

Due to our assumptions the sets \(L^2(X,m_t)\) and \(\text{Dom}(Ch_t)\) do not depend on \(t\). The domain \(\text{Dom}(Ch)\) of Cheeger’s energy endowed with the norm
\[
\sqrt{||u||^2_{L^2(X,m_t)} + Ch_t(u)}
\]
is a Banach space, cf. [7] Theorem 2.7.
In the following we additionally impose that for each $t$ the space $(X,d_t,m_t)$ is infinitesimally Hilbertian, i.e. Cheeger’s energy $C_t$ defines a quadratic form. In particular the domain is a separable Hilbert space and Lipschitz functions are dense, see [4]. In this case we will denote by $\mathcal{E}_t$ the associated Dirichlet form, which is the unique bilinear symmetric form satisfying

$$\mathcal{E}_t(u,u) = 2C_t(u) \quad \forall u \in \text{Dom}(C).$$

Moreover $\mathcal{E}_t$ is strongly local [4, Proposition 4.14], i.e.

$$u,v \in \text{Dom}(C), v \text{ constant on } \{u \neq 0\} \implies \mathcal{E}_t(u,v) = 0,$$

and admits the integral representation

$$\mathcal{E}_t(u,v) = \int_X \nabla_t u \cdot \nabla_t v \, dm_t \quad u,v \in \text{Dom}(C),$$

where

$$\nabla_t u \cdot \nabla_t v := \lim_{\varepsilon \to 0} \frac{|\nabla_t (u + \varepsilon v)|^2 - |\nabla_t u|^2}{2\varepsilon}$$

and the limit is understood in $L^1(X,m_t)$, see [4, Proposition 4.14].

We define the Laplace operator $\Delta_t$ as the generator of $\mathcal{E}_t$, i.e. as the unique non-positive self adjoint operator on $L^2(X,m_t)$ with domain $\text{Dom}(\Delta_t) \subset \text{Dom}(C)$ and

$$-\int_X \Delta_t uw \, dm_t = \mathcal{E}_t(u,v) \quad \forall u \in \text{Dom}(\Delta_t), v \in \text{Dom}(C).$$

We set $\mathcal{F} = \text{Dom}(C)$ endowed with $||u||^2_\mathcal{F} := ||u||^2_{L^2(X,m_t)} + \mathcal{E}_t(u)$ and $\mathcal{H} = L^2(X,m_t)$. We identify $\mathcal{H}$ with its own dual; the dual of $\mathcal{F}$ is denoted by $\mathcal{F}^*$. In particular we have $\mathcal{F} \subset \mathcal{H} \subset \mathcal{F}^*$ with dense and continuous embeddings. We define for $0 \leq s < \tau \leq T$ the Hilbert space

$$\mathcal{F}_{(s,\tau)} = L^2((s,\tau) \to \mathcal{F}) \cap H^1((s,\tau) \to \mathcal{F}^*),$$

equipped with the norm $(\int_\tau^\tau ||u_t||^2_\mathcal{F} + ||\partial_s u_t||^2_\mathcal{F} \, dt)^{1/2}$. According to Lemma 10.3 in [20] we have $\mathcal{F}_{(s,\tau)} \subset C([s,\tau] \to \mathcal{H})$.

**Definition 6.2.** A function $u$ is called solution to the heat equation

$$\partial_t u = \Delta_t u \text{ on } (s,\tau) \times X$$

if $u \in \mathcal{F}_{(s,\tau)}$ and if for all $w \in \mathcal{F}_{(s,\tau)}$

$$-\int_s^\tau \mathcal{E}_t(u_t,w_t) \, dt = \int_s^\tau \langle \partial_t u_t, w_t e^{-f_t} \rangle_{\mathcal{F}^*,\mathcal{F}} \, dt,$$

where $\langle \cdot, \cdot \rangle_{\mathcal{F}^*,\mathcal{F}}$ denotes the dual pairing.

A function $v$ is called solution to the adjoint heat equation

$$-\Delta_s v + \partial_s f \cdot v = \partial_s v \quad \text{on } (\sigma,\tau) \times X$$

if $v \in \mathcal{F}_{(\sigma,\tau)}$ and if for all $w \in \mathcal{F}_{(\sigma,\tau)}$

$$\int_\sigma^\tau \mathcal{E}_s(v_s,w_s) \, ds + \int_\sigma^\tau \int_X v_s \cdot w_s \cdot \partial_s f_s \, dm_s \, ds = \int_\sigma^\tau \langle \partial_s v_s, w_s e^{-f_s} \rangle_{\mathcal{F}^*,\mathcal{F}} \, ds.$$
By virtue of Theorem 2.2 and Theorem 2.5 in [15] we have existence and uniqueness to solutions of the heat and the adjoint heat equation with initial condition \( u_s = h \in \mathcal{H} \) and terminal condition \( v_t = h \in \mathcal{H} \) respectively. We denote these solutions by 
\[
    u_t(x) = P_{t,s} h(x), \quad v_s(x) = P^*_{t,s} h(x).
\]

Both solutions, called heat flow and adjoint heat flow respectively, satisfy
\[
    P_{t,s} h(x) = P_{t,r} \circ P_{r,s} h(x),
\]
\[
    P^*_{t,s} h(y) = P^*_{r,s} \circ P^*_{t,r} h(x).
\]
The operators are dual to each other in the sense that
\[
    \int (P_{t,s} u) v dm_t = \int u P^*_{t,s} v dm_s,
\]
and
\[
    \text{if } h \in \mathcal{H} \text{ with } 0 \leq h \leq 1 \text{ then } 0 \leq P_{t,s} h \leq 1
\]
\[
    P_{t,s} 1 = 1 \quad \text{whenever } m(X) < \infty, \quad (79)
\]
cf. Section 2 in [15].

### 6.1 Identification of the forward adjoint heat flow with the dynamic EDI-gradient flow for the entropy

We consider the adjoint heat flow \((\rho_t)_{0 \leq t \leq T}\) parametrized forwards in time, i.e. solving
\[
    \partial_t \rho_t = \Delta_t \rho_t + \rho_t \partial_t f_t \quad \text{on } (0, T) \times X
\]
with nonnegative initial data \( \rho_0 = h \). In the following we show coincidence of \((\rho_t)\) with the dynamic EDI-gradient flow \((\mu_t)\) of \( S \) via \( \mu_t = \rho_t m_t \). For this we have to assume that each \((X, d_t, m_t)\) satisfies CD(K, \( \infty \)). Following the approach in [3], we prove that \( \mu_t = \rho_t m_t \) is a dynamic EDI-gradient flow of \( S \). From the uniqueness it follows that both flows coincide.

**Lemma 6.3.** Let \( h \in \mathcal{H} \) and \((\rho_t)\) be the solution to the forward adjoint heat flow on \((0, T) \times X\) with \( \rho_0 = h \).

1. The flow \((\rho_t)\) is mass preserving, i.e.
\[
    \int \rho_t dm_t = \int h dm_0 \quad \forall 0 \leq t \leq T. \quad (80)
\]
2. If \( e: \mathbb{R} \to [0, \infty] \) is a convex lower semicontinuous function and \( e' \) is locally Lipschitz in \( \mathbb{R} \), it holds for a.e. \( t \in (0, T) \)
\[
    \frac{d}{dt} \int e(\rho_t) dm_t = -\int e''(\rho_t)|\nabla_t \rho_t|^2 dm_t + \int \partial_t f_t (\rho_t e'(\rho_t) - e(\rho_t)) dm_t. \quad (81)
\]
Proof. Since the measure is finite, $1 \in \mathcal{H}$, and hence by duality
\[ \int \rho_t \, dm_t = \int h \, dm_0. \]

In order to prove (81) we assume by a standard approximation that $e'$ is bounded and globally Lipschitz, cf. [3] Theorem 4.16. Since $e$ is convex and $\rho \in \mathcal{F}_{0,T}$ we have for $t_0 < t_1$
\[
\int e(\rho_{t_1}) \, dm_{t_1} - \int e(\rho_{t_0}) \, dm_{t_0} \\
\geq \int e'(\rho_{t_0})(\rho_{t_1} - \rho_{t_0}) \, dm_{t_1} + \int e(\rho_{t_0}) \, d(m_{t_1} - m_{t_0}) \\
= \int_{t_0}^{t_1} (\partial_t \rho_t, e'(\rho_t)e^{-f_t}) \, \, \mathcal{F}^{*} \, dt - \int_{t_0}^{t_1} \int e(\rho_t) \partial_t f_t \, dm_t \, dt \\
\geq \int_{t_0}^{t_1} \left( -\frac{1}{2} \left| \partial_t \rho_t \right|^2 \right) \, dt - \frac{1}{2} \left| e'(\rho_{t_0})e^{-f_{t_0}} \right|^2 - \int e(\rho_{t_0}) \partial_t f_t \, dm_t \right) \, dt,
\]
which is integrable. Changing the roles of $t_0$ and $t_1$ shows that $s \mapsto \int e(\rho_t) \, dm_t$ is absolutely continuous. Then, since $\rho \in \mathcal{F}_{0,T}$, we deduce from the mean value theorem for a.e. $t$
\[
\lim_{h \to 0} \frac{1}{h} \left( \int e(\rho_{t+h}) \, dm_{t+h} - \int e(\rho_t) \, dm_t \right) \\
= \lim_{h \to 0} \frac{1}{h} \int (e(\rho_{t+h}) - e(\rho_t)) e^{-f_{t+h}} \, dm + \lim_{h \to 0} \frac{1}{h} \int e(\rho_t)(e^{f_{t+h}} - e^{-f_t}) \, dm \\
= \lim_{h \to 0} \int e'(\rho_t) \frac{\rho_{t+h} - \rho_t}{h} \, dm_t - \int e(\rho_t) \partial_t f_t \, dm_t \\
= (\partial_t \rho_t, e'(\rho_t)e^{-f_t}) \, \mathcal{F}^{*} \, - \int e(\rho_t) \partial_t f_t \, dm_t,
\]

cf. [16] Corollary 5.5, [5] Lemma 12.3. Since $\rho$ is a solution to the forward adjoint heat equation we have
\[
\langle \partial_s \rho_s, e'(\rho_s)e^{-f_s} \rangle = -\mathcal{E}_s(\rho_s, e'(\rho_s)) + \int \rho_s e'(\rho_s) \partial_s f_s \, dm_s \\
= -\int e''(\rho_s) |\nabla_s \rho_s|^2 \, dm_s + \int \rho_s e'(\rho_s) \partial_s f_s \, dm_s,
\]
which proves (81).

Proposition 6.4. Let $(\rho_t)_{0 \leq t \leq T}$ be the solution of the forward adjoint heat equation with nonnegative initial datum $h \in \mathcal{H}$. Then it holds
\[
\int_0^t \int_{\{\rho_r > 0\}} \frac{|\nabla \rho_r|^2}{\rho_r} \, dm_r \, dr \leq \int h \log h \, dm_0 + \int h \, dm_0 \\
- m_t(X) + \int_0^t \int (\partial_r f_r) \rho_r \, dm_r \, dr,
\]
and the map $t \mapsto \int \rho_t \log \rho_t \, dm_t$ is locally absolutely continuous in $[0, T]$ and
\[
\frac{d}{dt} \int \rho_t \log \rho_t \, dm_t = -\int_{\{\rho_r > 0\}} \frac{|\nabla \rho_r|^2}{\rho_r} \, dm_t + \int (\partial_t f_t) \rho_t \, dm_t
\]
47
for a.e. $t \in [0, T]$.

Proof. By duality and (79), we have $\rho_t \geq 0$ for every $t \in (0, T)$. Applying formula (81) to $\rho_t + \varepsilon$ we get

$$
\frac{d}{dt} \int (\rho_t + \varepsilon) \log(\rho_t + \varepsilon) \, dm_t = -\int \frac{\nabla \rho_t}{\rho_t} \, dm_t + \int \partial_t f_t(\rho_t + \varepsilon) \, dm_t.
$$

Integrating from 0 to $t$ and letting $\varepsilon$ go to 0, we obtain by applying dominated and monotone convergence

$$
\int \rho_t \log \rho_t \, dm_t - \int \rho_0 \log \rho_0 \, dm_0 = \int_0^t \left( \int_{\{\rho_r > 0\}} \frac{\nabla \rho_r}{\rho_r} \, dm_r \right) \, dr.
$$

Using $\rho \log \rho \geq \rho - 1$ and the conservation of total mass (80) leads to

$$
\int_0^t \int_{\{\rho_r > 0\}} \frac{\nabla \rho_r}{\rho_r} \, dm_r \, dr \leq \int h \log h \, dm_0 + \int h \, dm_0 - m_t(X) + \int_0^t \left( \int_{\partial_r f_r} \rho_r \, dm_r \right) \, dr,
$$

which proves (82). As a consequence from (82) and (84) we get the local absolute continuity of $s \mapsto \int \rho_s \log \rho_s \, dm_s$ and (83).

The following two lemmas are crucial to conclude that the forward adjoint heat flow defines the EDE-gradient flow for the relative entropy. The first lemma gives an estimate of the squared slope of the entropy in terms of the Fisher information, which is an estimate in the static setting, while the second lemma represents a dynamic version of Kuwada’s Lemma, see e.g. [12, Proposition 3.7]. The proof of Proposition 6.6 relies on the dual formula of the dynamic distance $W_{t,s}$ (recall Definition 3.8) in terms of subsolutions to a modified Hamilton-Jacobi equation, cf. [15, Section 6].

**Proposition 6.5.** Assume $(X, d_t, m_t)$ satisfies CD($K$, $\infty$). For $\mu = \rho m_t \in \text{Dom}(S)$

$$
|\nabla S|_t^2(\mu) \leq \int_{\{\rho > 0\}} \frac{\nabla \rho_r}{\rho_r} \, dm_r.
$$

Proof. This is due to Theorem 9.3 in [3].

**Proposition 6.6.** Let $(\rho_t)_{0 \leq t \leq T}$ be the solution to the forward adjoint heat equation with nonnegative initial datum $h \in H$ such that $\int h \, dm_0 = 1$. Then the curve $t \mapsto \mu_t := \rho_t m_t$ is locally absolutely continuous and satisfies

$$
|\dot{\mu}_t|^2 \leq \int_{\{\rho_t > 0\}} \frac{\nabla \rho_t}{\rho_t} \, dm_t \quad \text{for a.e. } t \in [0, T].
$$

Proof. From (80) we know that $\int \rho_t \, dm_t = 1$ for every $0 \leq t \leq T$. Hence each $\mu_t = \rho_t m_t$ is a probability measure.
Let \( s < t \) and set \( \delta := t - s \). Then, with \( \vartheta(a) = s + a \), we define \( \text{HLS}_\vartheta \) as in \cite{15} Section 6] by

\[
\text{HLS}_\vartheta := \left\{ \varphi \in \text{Lip}_b([0, \delta] \times X) \left| \partial_a \varphi_a \leq -\frac{1}{2} |\nabla \vartheta(a)(\varphi_a)|^2 \right. \right\},
\]

and

\[
\tilde{W}^2_\vartheta(\mu_s, \mu_t) := 2 \sup_{\varphi} \left\{ \int \varphi \, d\mu_t - \int \varphi \, d\mu_s \right\},
\]

where the supremum runs over all maps \( \varphi(a, x) = \varphi_a(x) \in \text{HLS}_\vartheta \). Then we have by Lemma 6.5 in \cite{15}

\[
W^2_s(\mu_s, \mu_t) \leq e^{2L\delta} \tilde{W}^2_\vartheta(\mu_s, \mu_t).
\]

By applying \cite{2} Lemma 4.3.4] to the function \((a, b) \mapsto \int \rho_a \varphi_a dm_s\), where \( \varphi \in \text{HLS}_\vartheta \), we obtain

\[
\int \varphi \, d\mu_t - \int \varphi \, d\mu_s = \int_0^\delta \partial_a \int \varphi_a \, d\mu_{s+a} \, da
\]

\[
\leq \int_0^\delta \int -\frac{1}{2} |\nabla^{s+a}(\varphi_a)|^2 \, d\mu_{s+a} - E_{s+a}(\mu_{s+a}, \varphi_a) \, da
\]

\[
\leq \int_0^\delta \int -\frac{1}{2} |\nabla^{s+a}(\varphi_a)|^2 \, d\mu_{s+a}
\]

\[
+ \int \frac{1}{2} |\nabla^{s+a}(\varphi_a)|^2 \, d\mu_{s+a} + \frac{1}{2} \int \frac{|\nabla^{s+a}(\mu_{s+a})|^2}{\rho_{s+a}} \, dm_{s+a} \, da
\]

\[
= \int_0^\delta \int \frac{|\nabla^{s+a}(\mu_{s+a})|^2}{\rho_{s+a}} \, dm_{s+a} \, da.
\]

Taking the supremum over all \( \varphi \)

\[
W^2_s(\mu_s, \mu_t) \leq e^{2L\delta} \int \int \frac{|\nabla^{s+a}(\mu_{s+a})|^2}{\rho_{s+a}} \, dm_{s+a} \, da.
\]

Dividing by \( \delta^2 \) and letting \( \delta \to 0 \) we conclude

\[
|\frac{d}{ds} \mu_s|^2 \leq \int \frac{|\nabla \rho_s|^2}{\rho_s} \, dm_s.
\]

Now we are ready to prove our main result of this section.

**Theorem 6.7.** Let \((X, d_t, m_t)_{t \in [0, T]}\) be a family of Polish spaces with complete geodesic distances \( d_t \) satisfying \((12)\) such that \( m_t = e^{-f_t} m \), where \( m \in \mathcal{P}(X) \) and \( f_t \) are bounded functions satisfying and \((17)\). Assume that each static space satisfies RCD(\(K, \infty\)) for some finite number \( K \in \mathbb{R} \). Let \( h \in \mathcal{H} \) nonnegative with \( \bar{\mu} = h m_0 \).
1. Let $\rho_t$ solve the forward adjoint heat equation starting from $h$, then $\mu_t = \rho_t m_t$ is the dynamic EDE-gradient flow for the relative entropy $S_t$ starting in $\bar{\mu}$.

2. Conversely, let $\mu_t$ be the dynamic EDE-gradient flow for $S_t$, then $\mu_t = \rho_t m_t$ and $\rho_t$ is the solution to the forward adjoint heat equation.

Proof. Proposition 6.4 applied the forward flow $\rho_t$ yields
\[
\frac{d}{dt} \int \rho_t \log \rho_t \, dm_t = - \int_{\{\rho_t > 0\}} \frac{|\nabla \rho_t|^2}{\rho_t} \, dm_t + \int (\partial_t \rho_t) \rho_t \, dm_t.
\]

Integrating from 0 to $t$ and using Proposition 6.6 and Proposition 6.5 we obtain
\[
S_t(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_t|^2 \, dr + \frac{1}{2} \int_0^t |\nabla \mu_t|^2(\mu_t) \, dr \leq S_0(\bar{\mu}) + \int_0^t (\partial_t S_t)(\mu_t) \, dr.
\]

Moreover, by virtue of Proposition 2.8 in [15], $(\mu_t)$ is contained in the sublevel set of the entropy and hence, similarly as in the proof of Theorem 4.9, we get for all $t$
\[
S_t(\mu_t) - S_0(\bar{\mu}) \geq \int_0^t (\partial_t S_t)(\mu_t) \, dr - \int_0^t |\dot{\mu}| \nabla S_t(\mu_t) \, dr.
\]

Thus we have
\[
S_t(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_t|^2 \, dr + \frac{1}{2} \int_0^t |\nabla \mu_t|^2(\mu_t) \, dr = S_0(\bar{\mu}) + \int_0^t (\partial_t S_t)(\mu_t) \, dr.
\]

To show the converse implication, let $\tilde{\mu}_t$ be the solution to the adjoint heat equation parametrized forwards in time. From the previous argumentation we know that $\tilde{\mu}_t = \tilde{\rho}_t m_t$ is a dynamic EDE-gradient flow of the entropy. From Theorem 4.9 there is at most one gradient flow starting from $\bar{\mu}$, hence $\tilde{\mu}_t = \mu_t$ for every $t \in [0, T]$.

Remark 6.8. Let us recall the complete picture of forward and backward equation described in Section 6. The heat equation (forward in time) induces the adjoint heat equation (backward in time) and vice versa. Then $\mu_s := \rho_s m_s$, where $\rho_s$ denotes the adjoint heat flow (backward in time) is an upward dynamic EDI-gradient flow in the sense that
\[
S_s(\mu_s) + \frac{1}{2} \int_s^T |\dot{\mu}_r|^2 \, dr + \frac{1}{2} \int_s^T |\nabla \mu_r|^2(\mu_r) \, dr = S_T(\mu_T) + \int_s^T (\partial_r S_r)(\mu_r) \, dr.
\]

Equivalently, and this is what we showed, if $\mu_t = \rho_t m_t$, where $\rho_t$ solves the adjoint heat equation forward in time, then $\mu_t$ solves
\[
S_t(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_t|^2 \, dr + \frac{1}{2} \int_0^t |\nabla \mu_t|^2(\mu_t) \, dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r) \, dr.
\]

But then the heat equation is a backward equation.
Remark 6.9. If each \((X, d_t, m_t)\) is supposed to be \(RCD(K,N)\) space for finite numbers \(K,N\) and \((X, d_t, m_t)_{t \in [0,T]}\) is a super-Ricci flow (see [24] [22]) then it is shown in [15] that \(\mu_s = \rho_t m_s\) is characterized as the unique backward \(EVI(-2L, \infty)\)-gradient flow of the relativ entropy, i.e.

\[
\frac{1}{2}\partial_s W_{s,t}(\mu_s, t^2) \bigg|_{s=t} + LW^2(\mu, \sigma) \geq S_t(\mu_t) - S_t(\sigma).
\]

Then Proposition [3.11] already implies that \(\mu_t = \rho_t m_t\) is a dynamic \(EDE\)-gradient flow.

6.2 Identification of the heat flow with the dynamic gradient flow for Cheeger’s energy

In the following let \((X, d_t, m_t)_{t \in [0,T]}\) be a family of Polish metric measure spaces. We suppose that \((d_t)\) satisfies [12] and \(m_t = e^{-t} m\), where \(m\) is a \(\sigma\)-finite Borel measure on \(X\) and \((f_t)\) are Borel functions satisfying

\[
|f_t(x) - f_s(y)| \leq L^*|t - s|.
\]

We consider Cheeger’s energy \(\text{Ch}_t : L^2(X, m_t) \to [0, \infty]\), defined by

\[
\text{Ch}_t(u) = \frac{1}{2} \int_X |\nabla_t u|^2 dm_t,
\]

where \(|\nabla_t u|\) denotes the minimal relaxed gradient of \(u\). Since \(L^2(X, m_t)\) is a separable Hilbert space and the assumptions on the energy functional from Section 5 are satisfied by \((\text{Ch}_t)\), we directly obtain existence of a gradient flow in the sense of Definition 5.1.

**Theorem 6.10.** Let \(\bar{u} \in \text{Dom}(\text{Ch})\). Then there exists a unique gradient flow for \(\text{Ch}\) starting in \(\bar{u}\), i.e., an absolutely continuous curve \((u_t)\), solving

\[
\partial_t u_t = -D_t^* \text{Ch}_t(u_t) \quad \text{for a.e. } t \in (0, T)
\]

and \(\lim_{t \to 0} u_t = \bar{u}\).

**Proof.** Obviously \(\text{Ch}_t \geq 0\) for every \(t \in [0,T]\). Moreover \(u \mapsto \text{Ch}_t(u)\) is convex and lower semicontinuous by Theorem 4.5 in [3]. From Lemma 6.1 and (85) we obtain

\[
|\text{Ch}_t(u) - \text{Ch}_s(u)| \leq | \int |\nabla_t u|^2_2 - |\nabla_s u|^2_2| dm_t + | \int |\nabla_t u|^2_2 d(m_t - m_s)|,
\]

\[
\leq 2L\sup_{t \leq s} \int |\nabla_s u|^2_2 dm_s + L^* e^{L^*-L}[t-s]|t-s| \int |\nabla_s u|^2_2 dm_s,
\]

\[
\leq 2L\sup_{t \leq s} e^{C[t-s]} \int |\nabla_s u|^2_2 dm_s + L^* e^{L^*-L}[t-s]|t-s| \int |\nabla_s u|^2_2 dm_s,
\]

\[
\leq (2L + L^*) e^{L^*-L}[t-s]|t-s| \text{Ch}_s(u).
\]

We get the result as a consequence of Theorem 5.5. \(\square\)
In the case when the underlying spaces are infinitesimally Hilbertian we have by virtue of Theorem 2.2 in [15] a unique solution to the heat equation \( \partial_t u_t = \Delta_t u_t \) on \((0, T) \times X\) for each initial condition. Note that the solution is a priori ‘only’ contained in \( F(0, T) \) and hence \( \Delta_t u_t \) is not an element in \( L^2(X, m) \), or even an element in the subdifferential \( D^-\text{Ch}_t(u_t) \), which is at most single-valued since each \( \text{Ch}_t \) is a quadratic form.

Fortunately, by Theorem 2.12 in [15] it turns out that at least for a.e. \( t \in (0, T) \) \( u_t \in \text{Dom}(\Delta_t) \), thus \( \Delta_t u_t \in \mathcal{H} \). Consequently we may identify the gradient flow for Cheeger’s energy with the heat flow. This is the statement of the following theorem, for which we give an alternative proof using the dynamic EVI-property. In particular the heat flow can be constructed by the minimizing movement scheme from Section 5.

**Theorem 6.11.** Let \( (X, d_t, m_t)_{t \in [0, T]} \) be a family of Polish spaces with complete geodesic distances \( d_t \) satisfying (12) such that \( m_t = e^{-f_t} m \), where \( m \in \mathcal{P}(X) \) and \( f_t \) are bounded functions satisfying and (47). Assume that each static space is infinitesimally Hilbertian. Let \( \tilde{u}_t \) be the solution to the heat equation \( \partial_t \tilde{u}_t = \Delta_t \tilde{u}_t \) on \((0, T) \times X\) starting in some \( \bar{u} \in \text{Dom}(\text{Ch}) \). Then \( \tilde{u}_t \) satisfies

\[
\partial_t \tilde{u}_t \in -D^-\text{Ch}_t(\tilde{u}_t) \quad \text{for a.e. } t \in (0, T),
\]

and can be constructed as the limit of a minimizing movement scheme. Conversely, let \( u_t \) be the solution of the gradient flow of Cheeger’s energy \( \text{Ch}_t \). Then \( u_t \) solves the heat equation

\[
\partial_t u_t = \Delta_t u_t \quad \text{on } (0, T) \times X.
\]

In particular \( u_t = \tilde{u}_t \) in \( L^2(X, m) \) for every \( t \geq 0 \).

**Proof.** Both flows satisfy the dynamic EVI\((-L/2, \infty)\) gradient flow inequality almost everywhere by virtue of Proposition 5.2 and Theorem 2.16 in [15]. Hence from the contraction estimate (29)

\[
||u_t - \tilde{u}_t||_t^2 \leq e^{7L(t-s)}||u_s - \tilde{u}_s||_s^2 \quad \text{for a.e. } t \geq s,
\]

we obtain

\[
||u_t - \tilde{u}_t||_t^2 \leq \lim_{s \to 0} e^{7L(t-s)}||u_s - \tilde{u}_s||_s = 0 \quad \text{for a.e. } t,
\]

and hence by continuity \( ||u_t - \tilde{u}_t||_t = 0 \) for every \( t \). \( \square \)
References

[1] Luigi Ambrosio and Nicola Gigli. A user’s guide to optimal transport. In *Modelling and optimisation of flows on networks*, pages 1–155. Springer, Heidelberg, 2013.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the Space of Probability Measures*. Birkhäuser, Basel, 2005.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2013.

[4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.

[5] Luigi Ambrosio, Andrea Mondino, and Giuseppe Savaré. Nonlinear diffusion equations and curvature conditions in metric spaces. *arXiv:1509.07273*, 2015.

[6] Vladimir I. Bogachev. *Measure Theory*, volume 1. Springer-Verlag, Berlin, 2007.

[7] Jeff Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[8] Matthias Erbar. The heat equation on manifolds as a gradient flow in the Wasserstein space. *Annales de l'I. H. P. Probabilités et Statistiques*, 46(1):1–23, 2010.

[9] Matthias Erbar, Kazumasa Kuwada, and Karl-Theodor Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner’s inequality on metric measure spaces. *Invent. Math.*, 201:993–1071, 2015.

[10] Lucas Ferreira and Julio Valencia-Guevara. Gradient flows of time-dependent functionals in metric spaces and applications for PDE. *arXiv:1509.04161v1*, 2015.

[11] Nicola Gigli. On the heat flow on metric measure spaces: existence, uniqueness and stability. *Calc. Var.*, 39(1-2):101–120, 2010.

[12] Nicola Gigli, Kazumasa Kuwada, and Shin-ichi Ohta. Heat flow on Alexandrov spaces. *Comm. Pure Appl. Math.*, 66(3):307–331, 2013.

[13] Robert Haslhofer and Aaron Naber. Weak solutions for the Ricci flow I. *arXiv:1504.00911*, 2015.

[14] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29:1–17, 1998.

[15] Eva Kopfer and Karl-Theodor Sturm. Heat flows on Time-dependent Metric Measure Spaces and Super-Ricci Flows. *arXiv:1611.02570*, 2017.
[16] Janna Lierl and Laurent Saloff-Coste. Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. *arXiv:1205.6493*, 2012.

[17] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, pages 903–991, 2009.

[18] Robert McCann and Peter Topping. Ricci flow, entropy and optimal transportation. *American Journal of Mathematics*, 132(3):711–730, 2010.

[19] Shin-ichi Ohta and Karl-Theodor Sturm. Heat flow on Finsler manifolds. *Comm. Pure Appl. Math.*, 62(11):1386–1433, 2009.

[20] Michael Renardy and Robert C. Rogers. *An introduction to partial differential equations*, volume 13. Springer-Verlag, New York, 2004.

[21] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 7:97–169, 2008.

[22] Riccarda Rossi and Giuseppe Savaré. Gradient flows of non convex functionals in Hilbert spaces. *ESAIM Control Optim. Calc. Var.*, 12:564–614, 2006.

[23] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 169(1):65–131, 2006.

[24] Karl-Theodor Sturm. Super Ricci flows for metric measure spaces. I. *arXiv:1603.02193*, 2016.

[25] Peter Topping. *Lectures on the Ricci flow*, volume 325. Cambridge University Press, 2006.

[26] Cédric Villani. *Optimal transport, old and new*. Springer-Verlag, Berlin, Heidelberg, 2009.

[27] Max-Konstantin von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Comm. Pure and Appl. Math.*, 58(7):923–940, 2005.