TIME-DEPENDENT GRADIENT CURVES ON CAT(0) SPACES

C. JUN

ABSTRACT. We prove existence and uniqueness of time-dependent gradient curves for time-dependent functions with a convexity property on CAT(0) spaces. As an application, we prove existence and uniqueness of continuous pursuit curves, where the evader can be represented by a convex set or we can chase the barycenter of multiple evaders.

1. INTRODUCTION

This paper studies time-dependent gradient curves of time-dependent, almost convex (precisely, \(\lambda\)-convex) functions. The time-dependent case poses new challenges that do not arise in the time-independent setting.

Time-independent gradient flows have been studied extensively on CAT(0) spaces by Mayer [20] and on related metric spaces by Ambrosio, Gigli and Savaré [5] [9]. Independently, it was studied in CAT(K) spaces geometrically by Lytchak [19]. In a space with curvature bounded from below, it was studied by Petrunin [23]. In this paper we have invented a method, using the properties proved by Mayer about the gradient flow for a time-independent function, to obtain suitable approximations of the time-dependent case by piecewise fixed-time gradient segments, and show convergence without repeating the very complicated Crandall-Liggett scheme.

After we introduce Mayer’s results and modify them slightly in Section 4, we will use time-independent gradient curves to generate discrete solutions for the time-dependent case in Section 5. In order to prove existence, uniqueness and convergence estimates for time-dependent gradient curves, we must formulate appropriate dependence in the time variable (see Examples 22 and 23).

A CAT(K) space is a complete metric space such that no triangle is fatter than the triangle with same edge lengths in the model space of constant curvature K (see Definition 1). Like Mayer, we work on CAT(0) spaces, called nonpositively curved spaces or NPC spaces in [20]. For \(K = 0\), we can define CAT(0) spaces by convexity of squared distance functions (see Proposition 8). CAT(K) spaces include many important examples. Among many references, we mention [7] and [8]. Examples of CAT(0) spaces include simply connected Riemannian manifolds with non-positive sectional curvature (possibly with boundary satisfying a certain condition [1]) and trees. Spheres, surfaces of revolution, closed Euclidean domains with smooth boundary supported by spheres [2] and finite-dimensional spherical polyhedra with the link condition of Gromov [10], as well as all CAT(0) spaces, are examples of CAT(K) spaces for \(K > 0\).

1991 Mathematics Subject Classification. 91A24, 49N75, 53C20.

Key words and phrases. CAT(0) geometry, Time-dependent gradient curves, Pursuit-evasion games.
Our results on time-dependent gradient flows in turn feed back to pursuit-evasion games. For example, given a point $x_0$, the gradient flow of $\text{dist}_{x_0}$ defined by $\text{dist}_{x_0}(x) := d(x, x_0)$ is the geodesic flow toward $x_0$ as center. If the point moves, we will get time-dependent gradient curves. For a curve $E = E(t)$, we will show that there are time-dependent gradient curves for the time-dependent function $\text{dist}_{E(t)}$ in Section 6. Those curves are called (simple) continuous pursuit curves and $E$ is called the evader. As we shall show, we may also allow multiple evaders or uncertainty in evader position.

With different domains and different strategies, pursuit-evasion games have been considered by many mathematicians, computer scientists and engineers. The problems are generated from robotics, control theory and computer simulations. Under a simple pursuit strategy, the main constraints on pursuit-evasion are the geometry and topology of playing domains. Almost always these have been two-dimensional Euclidean domains, or higher-dimensional convex Euclidean domains. Recently there have been results on surfaces of revolution [11], cones [21], and round spheres [17]. Finally, $\text{CAT}(K)$ spaces were studied as a natural setting by Alexander, Bishop and Ghrist, because pursuit-evasion requires neither smoothness nor being locally Euclidean [2] [3]. $\text{CAT}(K)$ spaces include all previously studied domains and are vastly more general than have been usual in the extensive pursuit-evasion literature.

Recently, continuous pursuit games were applied to show the non-existence of shy-coupled Brownian motions in many Euclidean domains [6].

Very recently (after completion our work on this paper) we discovered Kim and Masmoudi’s paper [15] which is close to our work. However, our paper covers cases not covered by [15]. In particular, in our paper, we show that we get time-dependent gradient curves, which are defined on $[t_0, \infty)$ without an assumption that our flow has speed bounded uniformly on the entire space or that our flow has linear speed growth, whereas in [15] they need to use a Lipschitz constant with uniformly bounded speed. In Example 32 we give a specific example that is covered by our result but not by [15].

1.1. Outline of paper. In Sections 2 and 3, we list several properties of $\text{CAT}(K)$ spaces, $\lambda$-convex functions and gradient vectors. In Section 4, we look at Mayer’s work showing existence and uniqueness of time-independent gradient curves. Suppose that we have two $\lambda$-convex functions $F_{t_1}$ and $F_{t_2}$ for fixed-time $t_1$ and $t_2$. Then in Theorem 25, we examine the distance between two fixed-time gradient curves issuing from the same point. In this theorem, we need the Hölder continuity of gradient vectors in the time variable. We give an example illustrating the failure of the extendibility of time-dependent gradient curves in the absence of such a condition. Section 5 contains the statement and the proof of our main Theorem 28 showing existence and uniqueness of time-dependent gradient curves. In Section 6, Theorem 28 is applied to pursuit problems. In Section 7, we see a property of time-independent gradient curves.

2. $\text{CAT}(K)$ spaces

2.1. Metric spaces. Let $(X, d)$ be a metric space. A curve $\gamma : I \to X$ is called a geodesic if for all $t, t' \in I$, $d(\gamma(t), \gamma(t')) = c|t - t'|$ where $c$ is a constant, the speed of the geodesic $\gamma$. 


[xy] denotes a unit-speed geodesic γ from x to y defined on [0, t], where γ(0) = x, γ(t) = y and t = d(x, y). Δxyz denotes the geodesic triangle of geodesics [xy], [xz] and [yz].

A metric space is a geodesic space if any two points are joined by a geodesic; and a C-geodesic space if any two points with distance < C are joined by a geodesic.

2.2. CAT(K) spaces. M_0 denotes the 2-dimensional, complete, simply-connected space of constant curvature K. Then M_0 = B^2, M_1 = S^2 and M_{-1} = H^2. Let d_K be the metric of M_0. D_K denotes the diameter of M_0. Thus, D_K = \frac{\pi}{\sqrt{K}} if K > 0 and D_K = \infty if K \leq 0.

A triangle Δ\tilde{x}_1\tilde{x}_2\tilde{x}_3 in M_K is called a comparison triangle for Δx_1x_2x_3 in X if d_K(\tilde{x}_i, \tilde{x}_j) = d(x_i, x_j) for i, j \in \{1, 2, 3\}. We write

\tilde{ Δ}x_1x_2x_3 = Δ\tilde{x}_1\tilde{x}_2\tilde{x}_3.

Definition 1. Let (X, d) be a metric space (not necessarily locally compact) and K be a real constant. A complete D_K-geodesic space X is a CAT(K) space if for any geodesic triangle Δxyz in M_K, we have

d(z_1, z_2) \leq d_K(\bar{z}_1, \bar{z}_2),

where z_i is any point on [xy] and \bar{z}_i is the point on [\bar{xy}] such that d_K(\bar{z}_i, \bar{z}_i) = d(x, z_i) for i \in \{1, 2\}.

Let us define the (Alexandrov) angle between two geodesics.

Definition 2. Let γ_1, γ_2 be two geodesics in X starting at x. The (Alexandrov) angle \angle_x(γ_1, γ_2) between γ_1 and γ_2 is given by

\angle_x(γ_1, γ_2) := \lim sup_{t_1, t_2 \to 0} \theta(t_1, t_2)

where \theta(t_1, t_2) is the angle of Δ\bar{x}\bar{y}\bar{y}_2 at \bar{x} for y_1 = γ_1(t_1).

Note that we can get \theta(t_1, t_2) with the law of cosines. If X is a CAT(K) space, then \theta(t_1, t_2) is non-increasing in both variables. So there exists \lim_{t \to 0} \theta(t, t) and it is equal to \angle_x(γ_1, γ_2). For Δxyz in X, the angle \angle yxz of Δxyz at x is the Alexandrov angle between [xy] and [xz].

We need Helly’s Theorem for general CAT(0) spaces.

Theorem 3 (Helly’s Theorem). [18, Prop. 5.2] (see also [4, B.1]) Let X be a CAT(0) space and \{Y_a\}_{a \in I} be an arbitrary collection of closed bounded convex subsets of X.

If every finite index array \{a_1, a_2, \ldots, a_n\} \subset I satisfies

\bigcap_{i=1}^n Y_{a_i} \neq \emptyset,

then

\bigcap_{a \in I} Y_a \neq \emptyset.
2.3. **Tangent spaces.** We can generate a definition of a direction since the triangle inequality holds for angles between three geodesics and \( \angle_x(\gamma, \gamma) = 0 \) where \( \gamma \) is a geodesic starting at \( x \in X \). Two geodesics \( \gamma_1 \) and \( \gamma_2 \) starting at \( x \in X \) have the **same direction at** \( x \) if \( \angle_x(\gamma_1, \gamma_2) = 0 \) and we denote this relation by \( \gamma_1 \sim \gamma_2 \). This is an equivalence relation on the set of geodesics starting at \( x \). Then this set of equivalence classes is a metric space with metric \( \angle_x \). Denote the equivalence class of \( \gamma \) by \( \lbrack \gamma \rbrack \). Now consider the intrinsic metric \( d \) induced from \( \angle_x \). Note that if \( d(\lbrack \gamma_1 \rbrack, \lbrack \gamma_2 \rbrack) \leq \pi \), then \( d(\lbrack \gamma_1 \rbrack, \lbrack \gamma_2 \rbrack) = \angle_x(\lbrack \gamma_1 \rbrack, \lbrack \gamma_2 \rbrack) \). The completion of this space with metric \( d \) is called the **space of directions at** \( x \) and is denoted by \( \Sigma_x \).

The Euclidean cone over \( \Sigma_x \) is called the **tangent cone** \( T_x \) at \( x \); the elements of \( T_x \) are pairs \( v = (\xi, r) \) where \( \xi \in \Sigma_x \), \( r \geq 0 \) is a real number. We call \( \xi \) the direction of \( v \) and \( r \) the length of \( v \). All the pairs \( (\xi, 0) \) are identified as \( o_x \) and \( o_x \) is called the **vertex** of \( T_x \). The **norm** on \( T_x \) is given by \( r = ||(\xi, r)|| \), that is, it is the distance from the vertex \( o_x \) and the angle between \( (\xi, r), (\eta, s) \in T_x \), when both \( r, s \neq 0 \), is the same as the angle between \( \xi, \eta \).

The **inner product** \( \langle v, w \rangle \) for \( v, w \in T_x \) is defined by \( ||v|| ||w|| \cos \theta \) where \( \theta \) is the angle between \( v \) and \( w \) if \( v \neq o_x \) and \( w \neq o_x \). Otherwise, define \( \langle v, w \rangle = 0 \).

**Theorem 4.** [22] If \( X \) is a CAT(\( K \)) space, then \( \Sigma_x \) is a CAT(1) space, and \( T_x \) is a CAT(0) space.

**Definition 5.** Let \( X \) be a CAT(\( K \)) space, and \( \gamma : I \to X \) be a rectifiable curve. For \( t, t' \in I \) such that \( t \leq t' \), let \( \xi_{t'} \) be the direction at \( \gamma(t) \) of \( [\gamma(t), \gamma(t')] \). The curve \( \gamma \) has a **right-side tangent vector** \( \gamma'(t+) = (\xi, r) \) at \( \gamma(t) \) if there exist

\[
\begin{align*}
  r & = \lim_{t' \to t+} \frac{d(\gamma(t'), \gamma(t))}{t' - t}, \\
  \xi & = \lim_{t' \to t+} \xi_{t'}. 
\end{align*}
\]

Let us see the First Variation Formula for CAT(\( K \)) spaces.

**Theorem 6.** [7, Page 185] Let \( X \) be a CAT(\( K \)) space. For a distance \( f(t) \) between unit-speed geodesics \( \gamma_1(t) \) and \( \gamma_2(t) \), if \( f(0) < D_K \), then

\[
f'(0) = -\cos \alpha_1 - \cos \alpha_2
\]

where \( \alpha_1 \) is the angle at \( \gamma_1(0) \) between the geodesic \( \gamma_1 \) and the geodesic \( [\gamma_1(0), \gamma_2(0)] \).

3. **Semi-convex functions and their gradient vectors**

Suppose \( (X, d) \) is a CAT(\( K \)) space. For \( x_0, x_1 \in X \) and \( 0 \leq t \leq 1 \), let \( x_t \) be the point on \( [x_0, x_1] \) such that \( td(x_0, x_1) = d(x_0, x_t) \) and \( (1 - t)d(x_0, x_1) = d(x_1, x_t) \).

**Definition 7.** For \( \lambda \in \mathbb{R} \), a function \( F : X \to \mathbb{R} \) is \( \lambda \)-**convex** if

\[
F(x_t) \leq (1 - t)F(x_0) + tF(x_1) - t(1 - t)\frac{\lambda d^2(x_0, x_1)}{2}.
\]

for any \( x_0, x_1 \in X \).

**Proposition 8.** A geodesic metric space \( X \) is CAT(\( K \)) if and only if for any \( y, x_0 \) and \( x_1 \in X \),

\[
d^2(y, x_t) \leq (1 - t)d^2(y, x_0) + td^2(y, x_1) - t(1 - t)d^2(x_0, x_1).
\]

Briefly, if and only if functions \( x \mapsto d^2(y, x) \) are 2-convex.
Lemma 9. [20, Lemma 1.3] Let $X$ be a CAT(0) space and $F : X \to \mathbb{R}$. If $F$ is convex and lower semi-continuous, then $F$ is bounded from below on bounded subsets of $X$. Furthermore, the infimum of $F$ on each nonempty bounded convex closed subset of $X$ is attained.

To define gradient vectors of $\lambda$-convex functions, we need differentials of $\lambda$-convex functions.

Definition 10. For $x \in X$ and a locally Lipschitz function $F : X \to \mathbb{R}$, a function $d_x F : T_x \to \mathbb{R}$ is called the differential of $F$ at $x$ if for any curve $\gamma$ such that $\gamma(0) = x$ and $\gamma'(0+) = 0$ is defined,

$$d_x F(\gamma'(0+)) = \lim_{t \to 0^+} \frac{F \circ \gamma(t) - F \circ \gamma(0)}{t}.$$

In [16, Lemma 2.4], Kleiner showed if a $\lambda$-convex function $F : X \to \mathbb{R}$ is $L$-Lipschitz on a CAT(0) space $X$, then for every $x \in X$, there is a unique $L$-Lipschitz function $d_x F : T_x \to \mathbb{R}$. Moreover, $d_x F$ is convex and homogeneous of degree 1.

Then we can define gradient vectors of $\lambda$-convex functions.

Definition 11. A tangent vector $v \in T_x$ is called the downward gradient vector of $F$ at $x$ if

1. $(d_x F)(w) \geq -\langle v, w \rangle$ for all $w \in T_x$, and
2. $(d_x F)(v) = -\langle v, v \rangle$.

We denote $v$ by $\nabla_x (-F)$.

So the geometric meaning of the (downward) gradient vector is that $F$ will be decreased fastest in the direction of this gradient and the length of the gradient vector is the rate at which $F$ decreases in that direction.

Theorem 12. Let $X$ be a CAT(0) space. If $F$ is locally Lipschitz and $\lambda$-convex on $X$, then for any point $x \in X$, there is a unique downward gradient vector $\nabla_x (-F) \in T_x$.

Proof. For uniqueness, if $v, v'$ are two distinct downward gradient vectors of $F$ at $x$, then

$$||v||^2 = -(d_x F)(v) \leq \langle v, v' \rangle,$$

$$||v'||^2 = -(d_x F)(v') \leq \langle v, v' \rangle.$$ 

Then these inequalities imply that $||v|| = 0$ if and only if $\langle v, v' \rangle = 0$, hence if and only if $||v'|| = 0$ by the inner product definition. It follows that $v = v' = o_x$. Otherwise if $||v|| > 0$ and $||v'|| > 0$, by the inner product definition, we have

$$||v||^2 \leq ||v|| ||v'|| \cos \theta, \quad ||v'||^2 \leq ||v|| ||v'|| \cos \theta,$$

where $\theta$ is the angle between $v$ and $v'$. Therefore

$$1 \leq \cos^2 \theta$$

since $||v|| \leq ||v'|| \cos \theta \leq ||v|| \cos^2 \theta$. Since $\cos \theta > 0$ because $0 < ||v|| \leq ||v'|| \cos \theta$, we obtain $\cos \theta = 1$ and $\theta = 0$. Thus $v = v'$.

For existence, first if $d_x F \geq 0$ then $\nabla_x (-F)$ is defined to be $o_x$. Otherwise, let $r = \inf_{\eta \in \Sigma_x} (d_x F)(\eta) < 0$

where $\Sigma_x$ is the direction space at $x$. Let $S_x$ be the unit ball $\{w \in T_x, ||w|| \leq 1\}$ of the CAT(0) space $T_x$. Since $d_x F$ is Lipschitz and convex on $T_x$, $d_x F$ attains its
infimum on the nonempty bounded convex closed subset $S_x$ by Lemma 9. Since $d_x F$ is homogeneous, $\inf_{S_x} d_x F = r$. So we have a minimum direction $\xi$ such that $d_x F(\xi) = r$. Then $\nu = (\xi, \eta)$ satisfies the definition of the downward gradient vector, as follows:

1. When $\xi$ is the minimum point of $d_x F$ on the closed ball $S_x$, the convexity of $d_x F$ gives the support inequality
   \[ d_x F(\eta) \geq d_x F(\xi) \cos(s) = r(\xi, \eta), \]
   where $\eta \in \Sigma_x$ and $s = d_x (\xi, \eta) < \pi$.

   From the support inequality the proof of defining property (1) for the gradient vector $\nu$ easily follows from the homogeneity of $d_x F$ and $\langle \cdot, \cdot \rangle$:
   
   $\nu \in \Sigma_x$ and $s < \pi$,
   \[ d_x F(\eta) \geq r(\xi, \eta) = -r(\xi, \eta) = -(\nu, \eta). \]

   When $d_x (\xi, \eta) = \pi$, then the geodesic from $\xi$ to $\eta$ goes through the origin $o_x$ and the inequality we want is $d_x F(\eta) + d_x F(\xi) \geq 0$, as follows from the convexity of $d_x F$ on that geodesic.

2. $d_x F(\nu) = |r| d_x F(\xi) = |r|r = -|r|^2 (\xi, \xi) = -(\nu, \nu).$

\[ \square \]

We have an important lemma about gradient vectors at different points.

**Lemma 13.** [19][23, Lemma 1.3.3] Let $X$ be a CAT(0) space and $F$ be locally Lipschitz and $\lambda$-convex. Then for any two different points $x$ and $y$,

\[ \langle \xi_1, \nabla_x (-F) \rangle + \langle \xi_2, \nabla_y (-F) \rangle \geq \lambda d(x, y) \]

where $\xi_1$ is the direction of $[xy]$ at $x$ and $\xi_2$ is the direction of $[yx]$ at $y$.

**Proof.** By definition, we have

\[ \langle \xi_1, \nabla_x (-F) \rangle \geq -(d_x F)(\xi_1) = - \lim_{t \rightarrow 0^+} \frac{F \circ \gamma(t) - F \circ \gamma(0)}{t}, \]

where $\gamma = [xy]$. Since $F$ is $\lambda$-convex, we get

\[ \langle \xi_1, \nabla_x (-F) \rangle \geq \frac{F(x) - F(\gamma(d(x, y))) + \lambda d(x, y)^2/2}{d(x, y)}. \]

Doing similarly for $\langle \xi_2, \nabla_y (-F) \rangle$ and adding two inequalities, it is proved. $\square$

**Definition 14.** Let a function $F : X \rightarrow R$ be locally Lipschitz and $\lambda$-convex on $X$. A locally Lipschitz curve $\gamma : I \rightarrow X$ is a gradient curve of $F$ if for all $t \in I$, there exists the right-side tangent vector $\gamma'(t^+)$ and it is equal to the downward gradient vector $\nabla_{\gamma(t)}(-F)$ at $\gamma(t)$.

4. DISTANCE BETWEEN TWO FIXED-TIME GRADIENT CURVES

First, we will see Mayer’s work showing existence and uniqueness of time-independent gradient curves. Then we are going to look at two results about the distance between two fixed-time gradient curves: Corollary 20 and Theorem 25. These results will be used to prove our main Theorem 28. To emphasize the two results, we name them Distance I and Distance II.
In this section, we assume that \((X, d)\) is a CAT(0) space. For a function \(F = F(t, x)\) on \(\mathbb{R} \times X\), let \(F_t\) be given by \(F_t(x) := F(t, x)\). Then \(F\) is \(\lambda\)-convex on \(X\) if the function \(F_t\) is \(\lambda\)-convex.

In order to get Distance II, we will need the Hölder continuity of gradient vectors \(\nabla_x (-F_t)\) in the time variable \(t\). This need is illustrated by Example 23 below.

Distance I is essentially a result of Mayer about time-independent gradient curves. First, we need to define a step-energy function of \(F\).

**Definition 15.** [20, Def. 3] Given an initial position \(x_0 \in X\), an initial time \(t_0 \in \mathbb{R}\) and a time gap \(h > 0\), the step-energy function \(E_{t_0, x_0, h} : X \to \mathbb{R}\) at \((t_0, x_0)\) is defined by

\[
x \mapsto F(t_0, x) + \frac{1}{2h}d^2(x_0, x).
\]

Since for sufficiently small \(h\), \(E_{t_0, x_0, h}\) is a convex function with bounded sublevel sets, we can get a minimum value of \(E\). Differently from Mayer’s setting, we need to consider time \(t_0\) as variable. Mayer does not use the term “\(\lambda\)-convex” and he defines his condition on the function \(F\) in terms of the parameter \(S = -\lambda/2\). He also restricts to the case \(S > 0\), while we do not make any restriction on the sign of \(\lambda\).

**Proposition 16.** [20, Th. 1.8] Suppose that \(F\) is \(\lambda\)-convex and locally Lipschitz on a CAT(0) space \(X\). Let \(h > 0\) and if \(\lambda < 0\), let \(h < -1/2\lambda\). Then \(E_{t_0, x_0, h}\) has a unique minimum point on \(X\).

**Definition 17.** We will denote the unique minimum point of the step-energy function \(E_{t_0, x_0, h}\) by

\[
J(t_0, x_0, h).
\]

The function from \(X\) to \(X\) given by \(x \mapsto J(t_0, x, h)\), which we will call the discrete flow function with time gap \(h\), was studied in detail by Mayer. We have the following result, extending [20, Lemma 1.12] to \(\lambda \geq 0\), a longer interval for \(h\), and giving a slightly smaller Lipschitz constant instead of \(1/\sqrt{1+2\lambda h}\).

**Lemma 18.** [20, Lemma 1.12] Let \(h\) be as in Proposition 16. Then for any \(t_0 \in \mathbb{R}\), the discrete flow function \(x \mapsto J(t_0, x, h)\) is \((1 + \lambda h)^{-1}\)-Lipschitz.

This lemma is one of the primary tools to show that discrete flow converges well when \(h\) goes to zero.

For any \(y \in X\), let

\[
A(t_0) = -\min(0, \liminf_{d(x, y) \to \infty} \frac{F(t_0, x)}{d^2(x, y)}),
\]

\[
I_A(t_0) = \begin{cases} (0, \infty) & \text{if } A(t_0) = 0, \\ (0, \frac{1}{1 + A(t_0)}) & \text{if } A(t_0) > 0. \end{cases}
\]

Note that \(A(t_0)\) is independent of \(y\) because of the triangle inequality.

Mayer obtains only weak gradient curves in his main theorem since he assume the weaker condition that \(F\) is semi-continuous. Tangent vectors and gradient vectors were not contained in the definition of the weak gradient curve. If we assume \(F\) is locally Lipschitz on \(X\), by Theorem 12, we have gradient vectors everywhere and Mayer’s result can be modified for a gradient curve as follows (see Definition 14). We will see this modification in Section 7.
Theorem 19 (Existence and Uniqueness of time-independent gradient curves). [20, Th. 1.13] Let $X$ be a CAT(0) space. Suppose that $F$ is $\lambda$-convex and locally Lipschitz on $X$. For an initial position $x_0$ and fixed-time $t_0$, let $x^n_{t_0}(s) = x_0$ and

$$x^n_i(s) = \{(t_0, x^n_{t_0-1}(s), s/n)$$

where $i = 1, \cdots, n$. Then there is a unique gradient curve $\mu : \overline{I}_{\lambda} \rightarrow X$ of the function $F(t_0, \cdot)$ defined by $\mu(s) := \lim_{n \rightarrow \infty} x^n_{t_0}(s)$ such that $\mu(0) = x_0$.

Note that for simple notation, we use $x^n_i(s)$ even though it is also dependent on $x_0$. For each $n$, there are $n + 1$ points of $x^n_i(s)$’s and $\lim_{n \rightarrow \infty} x^n_{t_0}(s)$ is the limit of the sequence $(x^n_i(s)|n \in \mathbb{N})$ where $x^n_{t_0}(s)$ is the last point of $x^n_i(s)$’s.

Corollary 20 (Distance I). [20, Th. 2.1] Let $\mu_{\mu}$ be the gradient curve of the function $x \mapsto F(t_0, x)$ where $\mu_{\mu}(0) = y$. Then

$$d(\mu_{x_1}(s), \mu_{x_2}(s)) \leq e^{-\lambda s} d(x_1, x_2).$$

In order to obtain Distance II, we are going to assume that the function $t \mapsto \nabla_x(-F_t)$ is Hölder continuous. To see why we will need such a condition when we turn to time-dependent gradient curves, let us look at the following examples.

Definition 21. Let $X$ be a CAT(0) space and $F : \mathbb{R} \times X \rightarrow \mathbb{R}$ be $\lambda$-convex. A locally Lipschitz curve $\sigma : \{t_0, T\} \rightarrow X$ is a time-dependent gradient curve of $F$ at $x_0$ and $t_0$ if $\sigma(t_0) = x_0$, there exists the right-side tangent vector $\sigma'(t+)$ for all $t \in [t_0, T)$ and it is equal to the downward gradient vector $\nabla_{\sigma'(t)}(-F_t)$ at $\sigma(t)$.

We start with a time-independent example.

Example 22. Let $X$ be the subset of Euclidean plane such that $x \geq 0$ and $y \geq 0$. Let $F(x, y) := -\min(x, y)$. Then the gradient vector $\nabla_{(x, y)}(-F)$ is $(1, 0)$ if $x < y$, $(0, 1)$ if $y < x$ or $(1/2, 1/2)$ if $x = y$. Because $X$ is a manifold with boundary, we can put the tangent bundle metric on the set of all tangent vectors at all points.

Here $(x, y) \mapsto \nabla_{(x, y)}(-F)$ has discontinuities at the points where $y = x$ because its length is $1/\sqrt{2}$ at those points and length 1 everywhere else.

Thus for any initial point $(x_0, 0)$ on $X$, we have a gradient curve $\gamma$ of the function $F$ given by

$$\gamma(s) = \begin{cases} (x_0, s) & \text{if } 0 \leq s \leq x_0, \\ \left(\frac{1}{2}(x_0 + s), \frac{1}{2}(x_0 + s)\right) & \text{if } s \geq x_0. \end{cases}$$

Now we give a time-dependent example of a convex function having no time-dependent gradient curves at some points. In this example, the singular locus of Example 22 is translated $t$ units to the right for each $t \geq 0$.

Example 23. Let $X$ be the subset of Euclidean plane such that $x \geq 0$ and $y \geq 0$. For $t \geq 0$, let $F(t, x, y) := -\min(x - t, y)$. Then $(t, x, y) \mapsto \nabla_{(x, y)}(-F_t)$ has discontinuities at the points where $y = x - t$. If a time-dependent gradient curve leads to one of these points, it terminates and cannot be continued as a gradient curve. A gradient curve starting above the diagonal $y = x$ never reaches one of these points and hence is defined for all $t$. Those starting below the diagonal terminate in finite time when they get half way to the diagonal. No gradient curve can start at the diagonal. See Figure 1.

Lemma 24. Let $(X, d)$ be a CAT(K) space and $(T_x, \rho)$ be the tangent cone of $X$ at $x \in X$. Suppose $v_1, v_2, v_3 \in T_x$ and $\|v_3\| = 1$. Then

$$|\langle v_3, v_1 \rangle - \langle v_3, v_2 \rangle| \leq \rho(v_1, v_2).$$
Proof. Let $\theta_{ij} = \angle(v_i, v_j)$. Without loss of generality, assume $\theta_{13} \leq \theta_{23}$.

In the Euclidean upper half-plane, set

$$w_3 = (1,0), \quad w_i = \|v_i\|(\cos \theta_{13}, \sin \theta_{13}), \quad i = 1,2.$$  

Since projection to the $x$-axis does not increase distance,

$$|\langle v_3, v_1 \rangle - \langle v_3, v_2 \rangle| = \|v_1\|\cos \theta_{13} - \|v_2\|\cos \theta_{23} \leq \|w_1 - w_2\|.$$ 

By the triangle inequality for angles, $\theta_{23} - \theta_{13} \leq \theta_{12}$. If $\theta_{12} < \pi$, then

$$\|w_1 - w_2\| \leq \rho(v_1, v_2),$$

since the righthand side may be obtained from the lefthand side by increasing the hinge angle $\angle(w_1, w_2)$ from $\theta_{23} - \theta_{13}$ to $\theta_{12}$. If $\theta_{12} = \pi$, then $\rho(v_1, v_2) = \|w_1\| + \|w_2\|$ and the inequality still holds. □

After we prove our main Theorem 28, we will give several important examples in Section 6 that satisfy the assumption of Hölder continuity.

To finish this section, let us see how this Hölder continuity works on two fixed-time gradient curves issuing from the same point.

**Theorem 25 (Distance II).** Let $(X, d)$ be a CAT(0) space and $(T_x, \rho)$ be the tangent cone of $X$ at $x \in X$. Given $x_0 \in X$ and $t_1, t_2 \in \mathbb{R}$, suppose a function $F: \mathbb{R} \times X \to \mathbb{R}$ satisfies

1) $F$ is locally Lipschitz on $X$,
2) $F$ is $\lambda$-convex on $X$,
3) $\exists B > 0$ and $\alpha > 0$ such that $\rho(\nabla_x(-F_{t_1}), \nabla_x(-F_{t_2})) \leq B|t_1 - t_2|^\alpha$ for any $x \in X$.

Let $\mu_i: I_{A(t_i)} \to X$ be the fixed-time $t_i$ gradient curve of the function $x \mapsto F(t_i, x)$ where $\mu_i(0) = x_0$ for $i = 1,2$. Then there is a positive constant $T$ such that

$$d(\mu_1(s), \mu_2(s)) \leq 2Bs|t_1 - t_2|^\alpha$$

for all $s$ in $\overline{I_A(t_1)} \cap \overline{I_A(t_2)} \cap [0, T]$.

Note $I_{A(t_i)}$ is an interval dependent on $F_{t_i}$, given by Mayer (see Theorem 19). $\overline{I_A}$ is a closure of $I_A$.

**Proof.** Let $f(s)$ be a distance $d(\mu_1(s), \mu_2(s))$ and let $\xi_1(s)$ be the unit vector of $[\mu_1(s) \mu_2(s)]$ at $\mu_1(s)$ and $\xi_2(s)$ be the unit vector of $[\mu_2(s) \mu_1(s)]$ at $\mu_2(s)$.
By first variation formula, we have
\[ f^+(s) = -\langle \xi_1(s), \nabla_{\mu_1(s)}(-F_{t_1}) \rangle - \langle \xi_2(s), \nabla_{\mu_2(s)}(-F_{t_2}) \rangle. \]
We have three vectors \( \nabla_{\mu_2(s)}(-F_{t_1}), \nabla_{\mu_2(s)}(-F_{t_2}) \) and \( \xi_2(s) \) in a tangent cone \( T_{\mu_2(s)} \). By Lemma 24 with three vectors of \( T_{\mu_2(s)} \), we get
\[ f^+(s) \leq -\langle \xi_1(s), \nabla_{\mu_1(s)}(-F_{t_1}) \rangle - \langle \xi_2(s), \nabla_{\mu_2(s)}(-F_{t_1}) \rangle + \rho(\nabla_{\mu_2(s)}(-F_{t_1}), \nabla_{\mu_2(s)}(-F_{t_2})). \]
By Lemma 13 with \( x = \mu_1(s) \) and \( y = \mu_2(s) \), we have
\[ \langle \xi_1(s), \nabla_{\mu_1(s)}(-F_{t_1}) \rangle + \langle \xi_2(s), \nabla_{\mu_2(s)}(-F_{t_1}) \rangle \geq \lambda f(s). \]
Finally, we get
\[ f^+(s) \leq -\lambda f(s) + \rho(\nabla_{\mu_2(s)}(-F_{t_1}), \nabla_{\mu_2(s)}(-F_{t_2})) \leq -\lambda f(s) + B|t_1 - t_2|^\alpha. \]
Since \( f(s) \) is locally Lipschitz, \( f^+(s) \) exists for almost each \( s \). Thus we need to solve this differential equation. If \( \lambda = 0 \), it is proved trivially. So we only need to consider the case \( \lambda \neq 0 \). Solving this,
\[ d(\mu_1(s), \mu_2(s)) = f(s) \leq B|t_1 - t_2|^\alpha(1 - e^{-\lambda s})/\lambda \]
for all \( s \geq 0 \). Note that \( y = \frac{n}{m} +Ce^{mx} \) is the solution of \( y' = my+n \).
If \( \lambda \geq 0 \), it is proved since \( 1 - e^{-\lambda s} \leq \lambda t \).
If \( \lambda < 0 \), let \( \lambda_1 = -\lambda \). Then
\[ (1 - e^{-\lambda s})/\lambda = (1 - e^{\lambda_1 s})/(-\lambda_1) \]
\[ = (e^{\lambda_1 s} - 1)/\lambda_1 \]
for all \( s \geq 0 \). Then it is proved because there is an positive constant \( T \) such that \( e^{\lambda_1 s} - 1 \leq 2\lambda_1 s \) for all \( s \) in \( [0, T] \). \( \square \)

5. Time-dependent gradient curves

Now we will show the existence and uniqueness of time-dependent gradient curves (see Definition 21). For this, we need the definition of the flow map of the function \( F_t \) when \( F \) is Lipschitz in \( t \).

**Definition 26.** Let \( X \) be a CAT(0) space and \( F \) be \( \lambda \)-convex and locally Lipschitz on \( X \). Suppose that \( F \) is \( \lambda \)-Lipschitz in \( t \). Fix \( t \) and let \( \mu_{x,t} : I_A \to X \) be the fixed-time \( t \) gradient curve of the function \( F_t \) with \( \mu_{x,t}(0) = x \). Since \( F \) is \( \lambda \)-Lipschitz in \( t \), \( I_A \) (defined in (4.1)) is independent of \( t \). Then from Theorem 19, the flow map \( \Phi \) can be defined by
\[ \Phi : \mathbb{R} \times \overline{I}_A \times X \to X \]
\[ (t, \ell, x) \mapsto \mu_{x,t}(\ell) \]
(5.1)

For fixed \( t \), Mayer showed the semigroup property of this flow map \( \Phi \), which will be used in the proof of Theorem 28.

**Theorem 27.** [20, Th. 2.5] For \( \ell_1, \ell_2 \geq 0 \) such that \( \ell_1 + \ell_2 \in I_A \),
\[ \Phi(t, \ell_1 + \ell_2, x) = \Phi(t, \ell_2, \Phi(t, \ell_1, x)). \]
Next, we construct a set of piecewise fixed-time gradient curves beginning at \(x_0\), and show that they converge to a continuous curve. Let \(t^n_i(s) = t_0 + is/2^n\) for fixed \(s\) and an integer \(i \geq 0\). Let \(p^n_0(s) = x_0\) and let the vertices of the piecewise fixed-time gradient curves be

\[ p^n_i(s) = \Phi(t^n_i(s), s/2^n, p^n_{i-1}(s)). \]

See Figure 2. Note that \(p^n_i(s)\) is dependent on \(x_0\) and \(t_0\). But we use \(p^n_i(s)\) to avoid complicated notation.

We define piecewise fixed-time gradient curves \(\gamma^n_s : [t_0, \infty) \to X\) with step size \(s/2^n\) such that \(\gamma^n_s(t_0) = x_0\) given for \(t^n_{i-1}(s) \leq t \leq t^n_i(s)\) by

\[ \gamma^n_s(t) := \Phi(t^n_i(s), t - t^n_{i-1}(s), p^n_{i-1}(s)). \]

Two points \(p^n_{i-1}(s)\) and \(p^n_i(s)\) are connected by the fixed-time \(t^n_{i-1}(s)\) gradient curve of \(F_{t_i-1}(s)\). Then \(p^n_i(s)\) is connected to \(p^n_{i+1}(s)\) by the fixed-time \(t^n_i(s)\) gradient curve of \(F_{t_i}(s)\). The fixed-time \(t^n_i(s)\) gradient curve flows for time \(s/2^n\) from \(p^n_i(s)\) for each \(i \geq 0\). In Figure 2, there are two piecewise fixed-time gradient curves of \(\gamma^n_s\) and \(\gamma^n_{s+1}\). Note \(t^n_i(s) = t^n_{i+1}(s)\).

**Theorem 28** (Existence and Uniqueness of time-dependent gradient curves). Let \((X, d)\) be a CAT(0) space and \((T_x, \rho)\) be the tangent cone of \(X\) at \(x \in X\). Given \(x_0 \in X\) and \(t_0 \in \mathbb{R}\), suppose a function \(F : \mathbb{R} \times X \to \mathbb{R}\) satisfies

1) \(F\) is locally Lipschitz on \(X\),
2) \(F\) is \(\lambda\)-convex on \(X\),
3) \(F\) is \(L\)-Lipschitz in \(t\),
4) \(\exists B > 0, B_0 > 0\) and \(\alpha > 0\) such that \(\rho(-\nabla_x(F_t), -\nabla_x(F_{t'})) \leq B |t - t'|^\alpha\) for any \(x \in X\), any \(t, t' \in \mathbb{R}\) such that \(|t - t'|^\alpha \leq B_0\).

Then there is a time-dependent gradient curve \(\sigma_{x_0, t_0}\) of the function \(F\) at \(x_0\) and \(t_0\) defined by \(\sigma_{x_0, t_0}(t_0 + s) := \lim_{n \to \infty} \gamma^n_s(t_0 + s)\) such that \(\sigma_{x_0, t_0}(t_0) = x_0\). Moreover, any time-dependent gradient curve of \(F\) at \(x_0\) and \(t_0\) coincides with \(\sigma_{x_0, t_0}\).

Note that \(\gamma^n_s(t_0 + s) = p^n_{2^n}(s)\) since \(p^n_{2^n}(s)\) is \((2^n + 1)\)-th point of \(\{p^n_i(s)\}_{i \geq 0}\).

**Proof.** Since \(F\) is \(L\)-Lipschitz in \(t\), \(I_A\) (defined in (4.1)) is independent of \(t\). For a large integer \(n\), assume that \(h = s/2^n \in I_A \cap [0, T] \cap [0, B_0]\) where \(T\) is from Theorem 25.

**Claim 1.** The sequence \(\{p^n_{2^n}(s)\}\) is Cauchy. We define the limit curve \(\sigma_{x_0, t_0}\) by

\[ \sigma_{x_0, t_0}(t_0 + s) := \lim_{n \to \infty} p^n_{2^n}(s). \]

By induction on \(i\), we will show that

\[ d(p^n_i(s), p^n_{i+1}(s)) \leq B e^{-\lambda_0(i-1)h}[h/2]^{1+\alpha} \]

where \(\lambda_0 = \min[\lambda, \lambda]\).

The start of the induction is trivial since \(x_0 = p^n_0(s) = p^n_{2^n}(s)\). First, assume that

\[ d(p^n_{i-1}(s), p^n_{i+1}(s)) \leq B(i - 1)e^{-\lambda_0(i-2)h}[h/2]^{1+\alpha}. \]

Let \(z_i\) be the point on the fixed-time \(t_0 + (i - 1)h\) gradient curve flowing for time \(h = s/2^n\) from \(p^n_{2^n-1}(s)\), i.e

\[ z_i := \Phi(t^n_{i-1}(s), s/2^n, p^n_{2^n-1}(s)) \]
(see Figure 2). Note also that \( p^n_1(s) = z_1 \).

Since \( z_1 = \Phi(t^n_{i-1}(s), s/2^n, p^{n+1}_{2i-2}(s)) \) and \( p^n_1(s) = \Phi(t^n_{i-1}(s), s/2^n, p^n_{i-1}(s)) \), by Corollary 20, we get

\[
d(p^n_1(s), p^{n+1}_{2i}(s)) \leq d(p^n_1(s), z_1) + d(z_1, p^{n+1}_{2i}(s))
\]

\[
\leq e^{-\lambda_0 h}d(p^n_{i-1}(s), p^{n+1}_{2i-1}(s)) + d(z_1, p^{n+1}_{2i}(s)).
\]

By Theorem 27, the semigroup property of the fixed-time gradient flows gives that \( z_1 \) is equal to the point flowing for time \( h/2 \) from \( p^{n+1}_{2i-1}(s) \) on the same fixed-time to \( + (i-1)h \) gradient curve, that is,

\[
z_1 = \Phi(t^n_{i-1}(s), h/2, p^{n+1}_{2i-1}(s)).
\]

Since \( z_i = \Phi(t^n_{i-1}(s), h/2, p^{n+1}_{2i-1}(s)) \) and \( p^n_{2i+1}(s) = \Phi(t^n_{2i-1}(s), h/2, p^{n+1}_{2i-1}(s)) \), we can apply Theorem 25 at \( p^{n+1}_{2i-1}(s) \) with \( h/2 \leq T \). Thus we have

\[
d(z_i, p^{n+1}_{2i}(s)) \leq 2B(h/2)|t^n_{2i-1} - t^n_{i-1}|^\alpha = 2B(h/2)^{1+\alpha}.
\]

By Equations (5.2), (5.3) and (5.4), we obtain

\[
d(p^n_1(s), p^{n+1}_{2i}(s)) \leq e^{-\lambda_0 h}2B(i-1)e^{-\lambda_0 (i-2)h}(h/2)^{1+\alpha} + 2B(h/2)^{1+\alpha}
\]

\[
= 2B[(i-1)e^{-\lambda_0 (i-1)h} + 1](h/2)^{1+\alpha}
\]

\[
\leq 2Be^{-\lambda_0 (i-1)h}(h/2)^{1+\alpha}.
\]

Letting \( i = 2^n \), we have

\[
d(p^n_{2^n}(s), p^{n+1}_{2^{n+1}}(s)) \leq Bse^{-\lambda_0 s(\frac{s}{2^n})}\alpha.
\]

For \( m \) and \( n \) such that \( m > n \), from Equation (5.5),

\[
d(p^n_{2^n}(s), p^m_{2^m}(s)) \leq d(p^n_{2^n}(s), p^{n+1}_{2^{n+1}}(s)) + \cdots + d(p^m_{2^m-1}(s), p^m_{2^m}(s))
\]

\[
\leq Bs^{1+\alpha}e^{-\lambda_0 s(\frac{s}{2^n})}\alpha + \cdots + (\frac{1}{2^n})^\alpha
\]

\[
\leq Bs^{1+\alpha}e^{-\lambda_0 s(\frac{1}{2^n})}\alpha - \frac{1}{1-7/2^n}.
\]

This means that we have the Cauchy sequence \( (p^n_{2^n}(s)) \).
Claim 2.

\[ d(\sigma_{x_0, t_0}(t_0 + s), p^n_{2\gamma}(s)) \leq B's^{1+\alpha}\frac{1}{2\pi\rho}e^{-\lambda_0 s} \]

where \( \lambda_0 = \min(0, \lambda) \) and \( B' = \frac{B}{1-1/2\pi} \).

As \( m \to \infty \) in Equation (5.6), we have

\[ d(\sigma_{x_0, t_0}(t_0 + s), p^n_{2\gamma}(s)) \leq B's^{1+\alpha}\frac{1}{2\pi\rho}e^{-\lambda_0 s}. \]

In next claim, for each fixed integer \( n \) and any number \( s_1 \) smaller than \( s \), we will deal with the piecewise fixed-time gradient curves \( \gamma^i_{s_1} \) with step size \( s_1/2^n \), which is less than step size \( s/2^n \) of \( \gamma^n_{s_1} \). Note that \( \gamma^n_{s_1} \) is defined in the beginning of this section.

Claim 3. For any fixed values \( s, s_1 \) and \( t_0 \), the two sequences of points \( \{\gamma^n_{s_1}(t_0 + s)\}_{n \geq 0} \) and \( \{\gamma^n_{s_1}(t_0 + s)\}_{n \geq 0} \) get close in the sense that

\[ \lim_{n \to \infty} d(\gamma^n_{s_1}(t_0 + s), \gamma^n_{s_1}(t_0 + s)) = 0. \]

Let \( h_1 := s_1/2^n \) and \( h := s/2^n \). Let \( s \) and \( s_1 \) be fixed and let \( n \) be fixed until we take \( n \to \infty \). When \( s_1 = s/2^m \), it is proved by Claim 2.

Let \( x_t(s_1) = \gamma^n_{s_1}(t_0 + th_1) \) for an integer \( t \geq 0 \). So \( x_t(s_1) \) is the vertex of the piecewise fixed-time gradient curves \( \gamma^n_{s_1} \) and

\[ x_t(s_1) = \Phi(t_0 + (t-1)h_1, h_1, x_{t-1}(s_1)). \]

Note that \( t^n_i(s) = t_0 + is/2^n \) for \( i = 1, \ldots, 2^n \). Since \( s_1 < s \), there is a constant \( k = k(i) \) such that

\[ t_0 + kh_1 \leq t^n_i(s) < t_0 + (k + 1)h_1. \]

(see Figure 3).

Note that if \( s_1 = s/2^{n+1} \), then Figure 3 becomes Figure 2. We want to get an upper bound of the distance between two points \( \gamma^n_{s_1}(t^n_i(s)) \) and \( \gamma^n_{s_1}(t^n_i(s)) \). Here \( \gamma^n_{s_1}(t^n_i(s)) \) is the vertex of \( \gamma^n_{s_1} \). But \( \gamma^n_{s_1}(t^n_i(s)) \) is not necessarily the vertex of \( \gamma^n_{s_1} \) and there are several vertices \( x \)'s between \( \gamma^n_{s_1}(t^n_i(s)) \) and \( \gamma^n_{s_1}(t^n_{i-1}(s)) \). Specifically, by induction on \( i \), we will show that

\[ d(\gamma^n_{s_1}(t^n_i(s)), \gamma^n_{s_1}(t^n_i(s))) \leq 2B([s/s_1] + 1)is/2^n e^{-\lambda_0 is/2^n}. \]

In order to get the upper bound of the distance \( d(\gamma^n_{s_1}(t^n_i(s)), \gamma^n_{s_1}(t^n_i(s))) \), we will consider two distances \( d(\gamma^n_{s_1}(t^n_i(s)), z_1) \) and \( d(\gamma^n_{s_1}(t^n_i(s)), z_1) \) where \( z_1 \) will be defined below (see Figure 3).

Step 1. For \( d(\gamma^n_{s_1}(t^n_i(s)), z_1) \), first look at the point \( \gamma^n_{s_1}(t^n_{i-1}(s)) \). Let \( t' = t_0 + kh_1 \). Then the point \( \gamma^n_{s_1}(t^n_{i-1}(s)) \) is on the fixed-time \( t' \) gradient segment from \( \gamma^n_{s_1}(t') \) to \( \gamma^n_{s_1}(t' + h_1) = \Phi(t', h_1, \gamma^n_{s_1}(t')). \)

Since the gradient segment from \( \gamma^n_{s_1}(t^n_{i-1}(s)) \) to \( \gamma^n_{s_1}(t^n_i(s)) \) is the fixed-time \( t' \) gradient curve from \( \gamma^n_{s_1}(t^n_{i-1}(s)) \) to \( \gamma^n_{s_1}(t^n_i(s)) \) has the fixed-time \( t^n_{i-1}(s) \) gradient curve from \( \gamma^n_{s_1}(t^n_{i-1}(s)) \) to \( \gamma^n_{s_1}(t^n_i(s)) \) (in order to use Corollary 20, we need a fixed-time \( t^n_{i-1}(s) \) gradient curve from \( \gamma^n_{s_1}(t^n_{i-1}(s)) \). So let \( z_1 \) be the point on this fixed-time \( t_{i-1}(s) \) gradient curve flowing for time \( s/2^n \) from \( \gamma^n_{s_1}(t^n_{i-1}(s)) \). i.e

\[ z_1 := \Phi(t_{i-1}(s), s/2^n, \gamma^n_{s_1}(t^n_{i-1}(s))). \]
Since $\gamma^n_s(t^n_i(s)) = \Phi(t^n_i(s), s/2^n, \gamma^n_s(t^n_{i-1}(s)))$ and $z_i = \Phi(t^n_{i-1}(s), s/2^n, \gamma^n_{s_1}(t^n_{i-1}(s)))$,
by Corollary 20,
\begin{equation}
    d(z_i, \gamma^n_s(t^n_i(s))) \leq e^{-\lambda_0 s/2^n} d(\gamma^n_{s_1}(t^n_{i-1}(s)), \gamma^n_s(t^n_{i-1}(s))).
\end{equation}

**Step 2.** For $d(\gamma^n_s(t^n_i(s)), z_i)$, let
\[ y_{j+1} := \Phi(t_0 + (k + j)h_1, i\phi - (k + j)h_1, x_{j+k}(s_1)) \]
(see Figure 3). For every $j$, the point $y_{j+1}$ is in time $t_i = t_0 + i\phi$ since its initial point $x_{j+k}(s_1) = y_1(t_0 + (k + j)h_1)$ is in time $t_0 + (k + j)h_1$ and it flows for time $i\phi - (k + j)h_1$. Note that $y_{j/(s_1/s_1)-k+1} = \gamma^n_s(t^n_i(s))$. Also note that if $t_0 + kh_1 = t_{i-1}$, then
\[ y_1 = z_i. \]
By Theorem 27, the semigroup property of the fixed-time gradient flows gives that
\[ y_j = \Phi(t_0 + (k + j - 1)h_1, i\phi - (k + j)h_1, x_{j+k}(s_1)). \]
For three points $x_{j+k}(s_1)$, $y_j$ and $y_{j+1}$ when $j \geq 1$, from Theorem 25 at $x_{j+k}(s_1)$ with $T = i\phi - (k + j)h_1 \leq h = s/2^n$ and the time difference $h_1 = s_1/2^n$, we have
\begin{equation}
    d(y_{j+1}, y_j) \leq 2Bs/2^n(s_1/2^n)^\alpha.
\end{equation}
By Theorem 27, the semigroup property of the fixed-time gradient flows gives that
\[ y_1 = \Phi(t_0 + kh_1, s/2^n, \gamma^n_{s_1}(t^n_{i-1}(s))). \]
For three points $\gamma^n_s(t^n_{i-1}(s))$, $y_1$ and $z_i$, since $z_i = \Phi(t^n_{i-1}(s), s/2^n, \gamma^n_s(t^n_{i-1}(s)))$, we can apply Theorem 25 at $\gamma^n_{s_1}(t^n_{i-1}(s))$ with $T = s/2^n$. Thus we get
\begin{equation}
    d(z_i, y_1) \leq 2Bs/2^n(s_1/2^n)^\alpha
\end{equation}
since the difference of two fixed-time between the fixed-time gradient from $\gamma^n_{s_1}(t^n_{i-1}(s))$ to $y_1$ and the fixed-time gradient from $\gamma^n_{s_1}(t^n_{i-1}(s))$ to $z_i$ is less than $s_1/2^n$. 
Equations (5.9) and (5.10) imply that
\[
\begin{aligned}
d(y^n(s), z_i) &\leq d(y^n(t^n_i(s)), y^n(t^n_{i+1}(s))) + \cdots + d(y^n(z_i, t^n_1(s)) + d(y_1, z_i) \\
&\leq 2B([s/s_1]+1)s/2^n(s_1/2^n)^\alpha.
\end{aligned}
\] (5.11)

**Step 3.** We are ready for an induction argument to get Equation (5.7). The case \(i = 1\) is given by (5.11) since \(z_1 = y^n(t^n_1(s))\). Now assume that
\[
d(y^n(t^n_{i-1}(s), y^n(t^n_i(s))) \leq 2B([s/s_1]+1)(i-1)s/2^n e^{-\lambda_0(s)^{i-1}/2^n(s_1/2^n)^\alpha}.
\]
Then from Equations (5.8) and (5.11), we get
\[
d(y^n(t^n_i(s), y^n(t^n_{i+1}(s))) \leq d(y^n(t^n_i(s), z_i) + d(z_i, y^n(t^n_{i+1}(s))) \\
\leq d(y^n(t^n_i(s), z_i) + 2B([s/s_1]+1)s/2^n(s_1/2^n)^\alpha \\
\leq e^{-\lambda_0(s)/2^n}d(y^n(t^n_{i-1}(s)), y^n(t^n_i(s))) + 2B([s/s_1]+1)s/2^n(s_1/2^n)^\alpha \\
\leq e^{-\lambda_0(s)/2^n}2B([s/s_1]+1)(i-1)s/2^n e^{-\lambda_0(s)^{i-1}/2^n(s_1/2^n)^\alpha} \\
+ 2B([s/s_1]+1)s/2^n(s_1/2^n)^\alpha \\
= 2B([s/s_1]+1)(i-1)s/2^n e^{-\lambda_0(s)/2^n + s/2^n}(s_1/2^n)^\alpha \\
\leq 2B([s/s_1]+1)s/2^n e^{-\lambda_0(s)/2^n + s/2^n}(s_1/2^n)^\alpha.
\]

Letting \(i = 2^n\), we have
\[
d(y^n(t^n_0(s) + s, y^n(t^n_{2^n}(s))) \leq 2B([s/s_1]+1)s/2^n e^{-\lambda_0(s)/2^n}(s_1/2^n)^\alpha.
\]
Then taking \(n \to \infty\), Claim 3 is proved.

**Claim 4.** Limit curves satisfy the semigroup property.
For a limit curve \(\sigma_{x_0,t_0}\) such that \(\sigma_{x_0,t_0}(t_0) = x_0\), let \(x_1\) be \(\sigma_{x_0,t_0}(t_0 + s)\) and \(x_2\) be \(\sigma_{x_0,t_0}(t_0 + s + s_2)\). Then
\[
x_2 = \sigma_{x_1,t_0+s}(t_0 + s + s_2)
\]
where \(\sigma_{x_1,t_0+s}\) is the limit curve such that \(\sigma_{x_1,t_0+s}(t_0 + s) = x_1\).

For the proof, we need a piecewise fixed-time gradient curves with step size \(h = s/2^n\) from \(x_0\) and \(t_0\) converging to the limit curve \(\sigma_{x_0,t_0}\) as \(n \to \infty\). We can assume that \(s_2/s\) is not an integer, i.e. \(s_2/s > [s_2/s]\). For fixed \(s\) and \(s_2\), let
\[
t_i^n(s) = t_0 + is/2^n
\]
for \(i = 1, \ldots, N(n)\) and
\[
t_N(n+1) = t_0 + s + s_2
\]
where \(N(n) := 2^n(1 + [s_2/s])\). Let \(q_0 = x_0\) and for \(i = 1, \ldots, N(n) + 1\),
\[
q_i = \Phi(t^n_{i-1}(s), t^n_i(s) - t^n_{i-1}(s), q_{i-1}).
\]
Note that Since \(t^n_{2^n}(s) = t_0 + s\), by Claim 1, \(x_1 = \lim_{n \to \infty} q_{2^n}\) and by Claim 3, \(x_2 = \lim_{n \to \infty} q_{N(n)+1}\). Thus we get piecewise fixed-time gradient curves passing through the \(q_i\) converging to the limit curve \(\sigma_{x_0,t_0}\) as \(n \to \infty\).

For the limit curve \(\sigma_{x_1,t_0+s}\), we need another piecewise fixed-time gradient curves with step size \(h = s/2^n\) from \(x_1\) and \(t_0 + s\) converging to the limit curve \(\sigma_{x_1,t_0+s}\) as \(n \to \infty\). Let \(\hat{p}_n = x_1\) and
\[
\hat{p}_i = \Phi(t^n_{i-1}(s), t^n_i(s) - t^n_{i-1}(s), \hat{p}_{i-1})
\]
for \(i = 2^n + 1, \ldots, N(n) + 1\).
By Corollary 20 (Distance I), we get
$$d(\hat{\nu}_{N(n)} + 1, q_{N(n)}) \leq e^{-\lambda s} d(\hat{\nu}_2, q_{2n}).$$
Since $\hat{\nu}_2 = x_1$ and $x_1 = \lim_{n \to \infty} q_{2n}$, we get
$$\lim_{n \to \infty} d(\hat{\nu}_{N(n)} + 1, q_{N(n)}) = 0.$$
Since $x_2 = \lim_{n \to \infty} q_{N(N(n)) + 1}$ and $\sigma_{x_1, t_0 + s}(t_0 + s + s_2) = \lim_{n \to \infty} \hat{p}_{N(n)} + 1$ by Claim 3, the semigroup property of the limit curve is proved.

Claim 5. Let $\tilde{\sigma}$ be a reparametrization of the limit curve given by $s \mapsto \tilde{\sigma}(s) := \sigma_{x_0, t_0}(t_0 + t + s)$. Then for $s \in I_A$,
$$d(\tilde{\sigma}(s), \tilde{\sigma}(0)) \leq B's^{1+\alpha} e^{-\lambda_0 s},$$
where $\mu = \mu(s)$ is the fixed-time gradient curve of the function $\sigma \mapsto F(t_0 + t, \sigma)$ with $\mu(0) = \tilde{\sigma}(0) = \sigma_{x_0, t_0}(t_0 + t)$ and $\lambda_0 = \min(0, \lambda)$.

Note that Claim 5 implies
$$\lim_{s \to 0^+} \frac{d(\tilde{\sigma}(s), \tilde{\sigma}(0))}{s} = \lim_{s \to 0^+} \frac{d(\mu(s), \mu(0))}{s} = ||\nabla \tilde{\sigma}(0)(-F_{t_0 + t})||.$$ The second equality comes from the definition of time-independent gradient curve.

First, we look at $x_0$ and show the case $t = 0$. Letting $n = 0$ in Claim 2, $p_{2n}$ becomes $p(s)$ since $p_{2n}$ is just the time-independent gradient curve of $x \mapsto F(t_0, x)$ when $n = 0$. Thus we have
$$d(\tilde{\sigma}(s), \mu(s)) \leq B's^{1+\alpha} e^{-\lambda_0 s}$$

since $\tilde{\sigma}(s) = \sigma_{x_0, t_0}(s_0 + s)$.

Second, when $t > 0$, for any point $\tilde{\sigma}(0) = \sigma_{x_0, t_0}(t_0 + t)$, we can do this calculation again since we prove the semigroup property in Claim 4. Then we get the same inequality. Claim 5 is proved.

Claim 6. $\sigma_{x_0, t_0}$ is a time-dependent gradient curve of $F$ at $x_0$ and $t_0$ such that $\sigma_{x_0, t_0}(t_0) = x_0$.

For this, we will show the right-side tangent vector of $\sigma_{x_0, t_0}$ at $\sigma_{x_0, t_0}(t_0 + t)$ exists and is equal to the downward gradient vector $\nabla \sigma_{x_0, t_0}(t_0 + t)(-F_{t_0 + t})$.

Let $\tilde{\sigma}$ and $\mu$ be as in the Claim 5. Then since $\nabla \tilde{\sigma}(s)(-F_{t_0 + t})$ is the tangent vector of $\mu$ at $\mu(0)$, we need to show that the distance between the direction of $\mu$ at $\tilde{\sigma}(0)$ and the direction of a geodesic $[\tilde{\sigma}(0), \tilde{\sigma}(s)]$ goes to zero as $s \to 0^+$. Indeed, we need to show that the angle
$$\lim_{s \to 0^+} \angle(\tilde{\sigma}(0), \tilde{\sigma}(s), \mu(s))$$
is zero. This means that the direction of $\sigma_{x_0, t_0}$ at $\sigma_{x_0, t_0}(t_0 + t)$ is equal to the direction of $\mu$ at $\mu(0)$.

When $||\nabla \tilde{\sigma}(0)(-F_{t_0 + t})||$ is zero, it is trivial. So suppose that $||\nabla \tilde{\sigma}(0)(-F_{t_0 + t})||$ is not zero. So there is a constant $C$ such that $||\nabla \tilde{\sigma}(0)(-F_{t_0 + t})|| > 2C$. For sufficiently small $s$, we get $d(\mu(s), \mu(0)) \geq sC$ and $d(\tilde{\sigma}(s), \tilde{\sigma}(0)) \geq sC$, where $\mu(0) = \tilde{\sigma}(0)$, since
$$\lim_{s \to 0^+} \frac{d(\mu(s), \mu(0))}{s} = ||\nabla \tilde{\sigma}(0)(-F_{t_0 + t})|| = \lim_{s \to 0^+} \frac{d(\tilde{\sigma}(s), \tilde{\sigma}(0))}{s}.$$ By Claim 5, this implies that $\angle(\tilde{\sigma}(0), \tilde{\sigma}(s), \mu(s))$ is less than the angle of a Euclidean triangle with two edge lengths $sC$, $sC$ and third edge length less than $B's^{1+\alpha} e^{-\lambda_0 s}$. 


Claim 4 in the proof of Theorem 28.

Proof. This is Claim 2 in the proof of Theorem 28.

Let \( x \) and \( \xi \) where \( \lambda = \angle \nabla \sigma(x) \). Then a downward gradient vector \( s \) satisfies the conditions of Theorem 28. In particular, \( \sigma_i(t_0) = x_i \) where \( i = 1, 2 \),

\[
d(\sigma_1(t_0 + s), \sigma_2(t_0 + s)) \leq e^{-\lambda s} d(x_1, x_2).
\]

Then clearly, this estimate implies that any two time-dependent gradient curves with same initial position and time are same. This yields uniqueness of \( \sigma_{x_0,t_0} \).

Let \( f(s) \) be \( d(\sigma_1(t_0 + s), \sigma_2(t_0 + s)) \). By Lemma 13 with \( x = \sigma_1(t_0 + s) \) and \( y = \sigma_2(t_0 + s) \), we have

\[
(\xi_1(s), \nabla_{\sigma_1(t_0+s)}(-F_{t_0+s})) + (\xi_2(s), \nabla_{\sigma_2(t_0+s)}(-F_{t_0+s})) \geq \lambda f(s)
\]

where \( \xi_1(s) \) is the direction of \( [\sigma_1(t_0 + s)\sigma_2(t_0 + s)] \) and \( \xi_2(s) \) is the direction of \( [\sigma_2(t_0 + s)\sigma_1(t_0 + s)] \). By first variation formula, we get

\[
f^+(s) \leq -\lambda f(s).
\]

Solving this, Claim 7 is proved. \( \square \)

From the proof of Theorem 28, we extract the following corollaries.

Corollary 29. For any time-dependent gradient curves \( \sigma_i \) of \( F \) at \( x_i \) and \( t_0 \) such that \( \sigma_i(t_0) = x_i \) where \( i = 1, 2 \),

\[
d(\sigma_1(t_0 + s), \sigma_2(t_0 + s)) \leq e^{-\lambda s} d(x_1, x_2).
\]

Proof. This is Claim 7 in the proof of Theorem 28. \( \square \)

Corollary 30. For any time-dependent gradient curve \( \sigma_{x_0,t_0} \) and its piecewise fixed-time gradient curves of \( p^n_0 \) from Theorem 28,

\[
d(\sigma_{x_0,t_0}(t_0 + s), p^n_{2n}) \leq B s^{1+\alpha} \frac{1}{2\pi} e^{-\lambda_0 s}
\]

where \( \lambda_0 = \min(0, \lambda) \).

Proof. This is Claim 2 in the proof of Theorem 28. \( \square \)

Corollary 31. For any time-dependent gradient curve \( \sigma_{x_0,t_0} \), let \( x_1 = \sigma_{x_0,t_0}(t_0 + s) \) and \( x_2 = \sigma_{x_0,t_0}(t_0 + s + s_2) \). Then

\[
x_2 = \sigma_{x_1,t_0+s}(t_0 + s + s_2).
\]

Proof. This is Claim 4 in the proof of Theorem 28. \( \square \)

Example 32. Let \( X = \mathbb{R} \) and \( F: \mathbb{R}_{\geq 0} \times X \rightarrow \mathbb{R} \) given by

\[
F(t, x) = -(t + 1)x.
\]

Then a downward gradient vector \( \nabla_x(-F_t) \) of \( F_t \) at \( x \) is \( (t + 1) \). This example satisfies the conditions of Theorem 28. In particular,

\[
\rho(\nabla_x(-F_t),\nabla_x(-F_{t'})) = |t - t'|.
\]

So we can get a time-dependent gradient curves \( \sigma_{x_0} : [0, \infty) \rightarrow X \) given by

\[
\sigma_{x_0}(t) = 1/2t^2 + t + x_0.
\]
Let us look at the gradient flow \( \Phi : \mathbb{R}_{\geq 0} \times X \times [0, 1] \to X \) of \( F \). It is given by

\[
\Phi(t, x_0, h) = (t + 1)h + x_0.
\]

Then a value \( \rho(x_0, t) := \sup_{h \neq k} \frac{d(\Phi(t, x_0, h), \Phi(t, x_0, k))}{|h - k|} \) is equal to \( t + 1 \). But since \( \rho(x_0, t) \) is not bounded in \( t \), \( \Phi \) does not have linear speed growth (see the following definition for linear speed growth). So we have a time-dependent gradient curve \( \sigma_{x_0} \) defined on \([0, \infty)\) even if \( \Phi \) doesn’t have linear speed growth.

**Definition 33.** Given a flow \( \Phi : \mathbb{R}_{\geq 0} \times X \times [0, 1] \to X \), and fixed \( t, r, l \) and \( x \), let

\[
\rho(x, t; r, l) := \sup_{d(x, y) < r, |t - s| \leq l} \rho(y, s)
\]

where \( \rho(x, t) := \sup_{h \neq k} \frac{d(\Phi(t, x, h), \Phi(t, x, k))}{|h - k|} \).

Then \( \Phi \) is said to have **linear speed growth** if there is a point \( x \in X \) with positive constants \( c_1(x) \) and \( c_2(x) \) such that for all \( r > 0 \) and \( t, l > 0 \),

\[
\rho(x, t; r, l) \leq c_1(x)r + c_2(x).
\]

6. **Application: Simple pursuit**

6.1. **Simple pursuit curves chasing convex sets.** Suppose \( (X, d) \) is a CAT(0) space. We use the Hausdorff metric \( d_H \) on the set of closed convex sets.

Let \( d_Y : X \to \mathbb{R} \) be the distance to a complete convex subset \( Y \subset X \), defined by

\[
x \mapsto d_Y(x) := \min_{y \in Y} d(y, x).
\]

**Proposition 34.** (See [7, Page 178]) Let \( Y \) be a complete convex subset. Then the distance function \( d_Y \) is convex and for \( x \in X \), there is a unique point \( x_0 \) of \( Y \) such that \( d_Y(x) = d(x_0, x) \).

This point \( x_0 \) is called the **footpoint of** \( x \) in \( Y \).

**Proposition 35.** Let \( Y_t \) be a curve of closed convex sets in \( X \) such that \( d_H(Y_t, Y_{t'}) \leq |t - t'| \) and \( q_t \) be the footpoint of \( p \) in \( Y_t \). Then we have

\[
|d(p, q_t) - d(p, q_{t'})| \leq |t - t'|
\]

and

\[
d(q_t, q_{t'}) \leq \left( \sqrt{|t - t'|^2 + 2|d(p, q_t) + |t - t'|} \right) \sqrt{|t - t'|}
\]

for any \( t \) and \( t' \).

**Proof.** Let \( x \) be the footpoint of \( q_t \) in \( Y_t \), and \( y \) be the footpoint of \( q_{t'} \) in \( Y_t \). For all \( \epsilon > 0 \),

\[
d(x, q_t) \leq |t - t'| + \epsilon
\]

and

\[
d(y, q_{t'}) \leq |t - t'| + \epsilon
\]

since \( d_H(Y_t, Y_{t'}) \leq |t - t'| \). Hence,

\[
(6.1) \quad d(x, q_t) \leq |t - t'| \quad \text{and} \quad d(y, q_{t'}) \leq |t - t'|.
\]

By the triangle inequality, we have

\[
(6.2) \quad d(p, x) - d(p, q_t) \leq |t - t'|
\]

and

\[
(6.3) \quad d(p, y) - d(p, q_{t'}) \leq |t - t'|.
\]
By the definition of $q_t'$, we get
\[ d(p, q_t') \leq d(p, x) \leq d(p, q_t) + |t - t'|. \]
Similarly, by the definition of $q_t$,
\[ d(p, q_t) \leq d(p, y) \leq d(p, q_{t'}) + |t - t'|. \]
This means
\[ (6.4) \quad \left| d(p, q_t) - d(p, q_{t'}) \right| \leq |t - t'|. \]
Let $z_s$ be the point on $[z_0 z_1]$ such that $s d(z_0, z_1) = d(z_0, z_s)$ and $(1-s) d(z_0, z_1) = d(z_1, z_s)$ for $0 \leq s \leq 1$. For $[q_t y]$, suppose that $z_0 = q_t$, and $z_1 = y$. Since $q_t$ and $y$ are in $Y_t$, $[q_t y]$ lies in $Y_t$. Then, by Proposition 8, we get
\[ d(p, q_t)^2 \leq d(p, z_s)^2 \leq (1-s) d(p, q_t)^2 + s d(p, y)^2 - s(1-s) d(q_t, y)^2. \]
This gives
\[ s(1-s) d(q_t, y)^2 \leq s d(p, y)^2 - s d(p, q_t)^2. \]
Dividing both sides by $s$ and taking $s \to 0$, we have
\[ d(q_t, y)^2 \leq d(p, y)^2 - d(p, q_t)^2. \]
Then it becomes
\[ d(q_t, y)^2 \leq (d(p, y) + d(p, q_t))(d(p, y) - d(p, q_t)). \]
With (6.3) and (6.4), it yields
\[ d(q_t, y)^2 \leq (2d(p, q_t) + 2|t - t'|)(2|t - t'|). \]
Then
\[ d(q_t, y) \leq 2 \sqrt{d(p, q_t)} + |t - t'| \sqrt{|t - t'|}. \]
Therefore by (6.1), we have
\[ d(q_t, q_{t'}) \leq |t - t'| + 2 \sqrt{d(p, q_t)} + |t - t'| \sqrt{|t - t'|}. \]

\[ \square \]

\textbf{Figure 4.} The footpoint $q_t$ of $p$ in $Y_t$ and the closed ball $B_{x_0}(L)$
Theorem 36 (α = 1/2). Let \( Y_t \) be a curve of closed convex sets in \( X \) such that \( t \in [0, \infty) \) and \( d_{Y_t}(Y_t, Y_{t'}) \leq |t - t'|. \) For any \( x_0 \in X, \) if there is a constant \( T > 0 \) such that \( \min_{t \in [0, T]} d_{Y_t}(x_0) > 2T, \) then inside the closed ball \( B_{x_0}(T), \) there is a unique time-dependent gradient curve \( \sigma_{x_0} : [0, T] \to X \) of the function \( F \) where \( F(t, p) := d_{Y_t}(p) \) such that \( \sigma_{x_0}(0) = x_0. \)

\( \sigma_{x_0} \) is called a simple pursuit curve chasing the convex set \( Y_t \) with an initial \( x_0. \)

Proof. If it exists, its speed is one. So it must be inside the ball \( B_{x_0}(T). \) For \( x, z \in X, \) let \( x_t \) be the footpoint of \( x \) in \( Y_t \) and \( z_t \) be the footpoint of \( z \) in \( Y_t. \) Suppose that \( d(x, x_t) \leq d(z, z_t). \) Let \( z' \) be the point on \( [zz_t] \) at distance \( d(x, x_t) \) from \( z_t. \) Since \( z' \) is the footpoint of \( z \) in \( \{ y \in X \mid d_{Y_t}(y) \leq d(x, x_t) \}, \)

\[
d(z', z) \leq d(x, z).
\]

Then we have

\[
F(t, z) - F(t, x) = d(z, z_t) - d(x, x_t) = d(z', z) \leq d(x, z).
\]

Thus \( F \) is 1-Lipschitz on \( X. \)

From Proposition 34, \( F \) is \( \lambda \)-convex on \( X \) with \( \lambda = 0. \) By Proposition 35, we know that \( F \) is \( L \)-Lipschitz in \( t \) with \( L = 1. \) Thus we see that \( F \) satisfies three conditions of Theorem 28.

Next, we will show that \( F \) satisfies last condition of Theorem 28. Let \( R = \max_{t \in [0, T]} d_{Y_t}(x_0) := \max_{t \in [0, T]} d_{Y_t}(x_0). \) Then \( R > 2T. \)

For \( t, t' \in [0, T], \) suppose that \( |t - t'| \leq R \) and \( 5 \sqrt{\sqrt{R}/|t-t'|} \) is smaller than \( T/2. \) For a point \( p \in B_{x_0}(T), \) let \( q_t \) be the footpoint of \( p \) in \( Y_t \) and \( q_{t'} \) be the footpoint of \( p \) in \( Y_{t'}. \)

Then by triangle inequality,

\[
d(p, q_t) + d(p, x_0) \geq d(x_0, q_t) \geq d(x_0, Y_t) > 2T.
\]

Since \( d(p, x_0) \leq T, \) this implies \( d(p, q_t) > T. \) Similarly, we get \( d(p, q_{t'}) > T. \)

Also \( d(p, q_t) \leq d(p, x_0) + d(x_0, Y_t) \leq T + R \leq 2R. \) Then from Proposition 35,

\[
d(q_t, q_{t'}) \leq (\sqrt{|t - t'|} + 2 \sqrt{\sqrt{R} + |t - t'|}) \sqrt{|t - t'|} \\
< 5 \sqrt{\sqrt{R}/|t-t'|}.
\]

Thus \( d(q_t, q_{t'}) \leq T/2. \)

If we consider a comparison triangle \( \triangle \tilde{q}_t \tilde{p} \tilde{q}_{t'}, \) since \( d(\tilde{p}, \tilde{q}_t) > T, \)

\[
\sin \angle \tilde{q}_t \tilde{p} \tilde{q}_{t'} \leq d(\tilde{q}_t, \tilde{q}_{t'})/T.
\]

Since \( d(\tilde{q}_t, \tilde{q}_{t'})/T = d(q_t, q_{t'})/T \leq 1/2, \) this implies

\[
\angle \tilde{q}_t \tilde{p} \tilde{q}_{t'} \leq \frac{2d(\tilde{q}_t, \tilde{q}_{t'})}{\sqrt{3}T}
\]

because \( \sin^{-1} \) has Lipschitz constant \( 2/\sqrt{3} \) on \( [0, 1/2]. \)

By Equations (6.5) and (6.6), we have

\[
\angle \tilde{q}_t \tilde{p} \tilde{q}_{t'} \leq \frac{10 \sqrt{R}}{\sqrt{3}C} \sqrt{|t - t'|}.
\]
Since \(\|\nabla_p(-F_t)\| = 1\), it implies
\[
\rho(\nabla_p(-F_t), \nabla_p(-F_{t'})) = \angle q_t p q_{t'} \leq \frac{10\sqrt{R}}{\sqrt{3}C} \sqrt{|t-t'|}.
\]

By Theorem 28 with \(\alpha = 1/2\), \(B = \frac{10\sqrt{R}}{\sqrt{3}C}\), \(B_0 = \min\{R, T^2/100R\}\) and \(X = B_{x_0}(T)\), we have a time-dependent gradient curve starting at \(x_0\).

**Theorem 37.** Let \(Y_t\) be a curve of closed convex sets in \(X\) such that \(t \in [0, \infty)\) and \(d_{t1}(Y_{t1}, Y_{t'}) \leq |t - t'|\). Then for any \(x_0 \in X\), there is a unique time-dependent gradient curve \(\sigma_{x_0}\) of the function \(F\) where \(F(t, p) := d_{Y_t}(p)\) such that \(\sigma_{x_0}(0) = x_0\). When it meets the curve of closed convex sets in \(X\) at time \(t_1\), that is, \(d(Y_{t1}, \sigma_{x_0}(t_1)) = 0\), it will stop.

**Proof.** For the initial point \(x_0\), let \(f(s) = \min_{t \in [0, s]} d_{Y_{t1}}(x_0)\). If there is a constant \(T_{x_0} > 0\) such that \(\min_{t \in [0, T_{x0}]} d_{Y_{t1}}(x_0) > 2T_{x0}\), by Theorem 36, then inside the closed ball \(B_{x_0}(T_{x0})\), there is a unique time-dependent gradient curve \(\sigma_{x_0} : [0, T_{x0}] \to X\) of the function \(F\) where \(F(t, p) := d_{Y_t}(p)\) such that \(\sigma_{x_0}(0) = x_0\).

If there is no such constant \(T_{x0}\), it means that \(d(x_0, Y_0) = 0\) since \(f(s_1) \geq f(s_2)\) when \(0 \leq s_1 \leq s_2\) and \(f\) is continuous. When it arrives at \(x_1 = \sigma_{x_0}(T_{x0})\), we find a constant \(T_{x1}\) again. If there is a constant \(T_{x1}\) for \(x_1\), by Theorem 36, we can extend our gradient curve at \(x_1\) or it stops at \(x_1\). This means that we will have an interval \([0, t_1]\) such that the time-dependent gradient curve \(\sigma_{x_0}\) exists on \([0, t_1]\) and \(d(\sigma_{x_0}(t_1), Y_{t1}) = 0\) or \(\sigma_{x_0}\) exists on infinite time. \(\square\)

For \(Y_t\), if \(Y_t\) is just a point of \(X\), then we have the following corollary.

**Corollary 38.** Let \(Y_t\) be a curve of points with speed \(\leq 1\). Then for any \(x_0 \in X\), there is a unique simple pursuit curve \(\sigma_{x_0}\) chasing the curve \(Y = Y_t = Y_t\) such that \(\sigma_{x_0}(0) = x_0\).

In [14], we study continuous pursuit curves chasing a moving point on \(\text{CAT}(K)\) spaces and geometrically show existence and uniqueness of continuous pursuit curves on \(\text{CAT}(K)\) spaces. Additionally, we get regularity of continuous pursuit curves which is a replacement for \(C^{1,1}\) regularity in smooth spaces.

6.2. **Simple pursuit curve chasing barycenters.** Suppose that \(X\) is a \(\text{CAT}(0)\) space. When we deal with multiple evaders in pursuit-evasion games, we still may find a strategy to get a continuous pursuit curve. For this, we consider the barycenter of multiple points.

Since for points \(x_1, \cdots, x_n \in X\), the function \(x \mapsto \sum_{i=1}^n d^2(x, x_i)\) is strictly convex, there exists a unique minimum point of this function.

**Definition 39.** [12, Def. 2.3] Given \(n\) points \(x_i\) on \(X\), the barycenter \(b\) of the \(x_i\)'s is defined to be the minimum point of the function \(x \mapsto \sum_{i=1}^n d^2(x, x_i)\).

Let \(\mathcal{P}(X)\) denote the set of all probability measures \(\nu\) on \((X, \mathcal{B}(X))\) with separable support \(\text{supp}(\nu) \subset X\) where \(\mathcal{B}(X)\) is the set of Borel sets of \(X\).

More generally,

**Definition 40.** [25, Prop. 4.3] For \(\nu \in \mathcal{P}(X)\) such that \(\int_X d^2(x, y)\nu(dy) < \infty\) for some (hence all) \(x \in X\), the minimum point of the function \(x \mapsto \int_X d^2(x, y)\nu(dy)\) is called the barycenter \(b(\nu)\) of \(\nu\).
Since the function \( x \mapsto \int_X d^2(x, y)\nu(dy) \) is continuous and strictly convex, the barycenter of \( \nu \) is well-defined.

Let \( \delta_x \) be the Dirac measure of \( x \) given by \( \delta_x(A) = 1 \) if \( x \in A \) or \( \delta_x(A) = 0 \) otherwise, for any subset \( A \) of \( X \).

**Lemma 41.** Set \( \nu = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \). Then the barycenter \( b(\nu) \) of \( \nu \) is equal to the barycenter of the \( x_i \)'s.

**Proof.** Since \( \nu = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \),

\[
\int_X d^2(x, y)\nu(dy) = \frac{1}{n} \sum_{i=1}^{n} d^2(x, x_i).
\]

Then the minimum point of the function \( x \mapsto \sum_{i=1}^{n} d^2(x, x_i) \) is equal to the barycenter of \( \nu \).

Here we want to give a strategy for chasing multiple evaders. Given \( n \) evaders \( E_i = E_i(t) \) with speed \( \leq 1 \) in \( X \), we have the barycenter curve \( b = b(t) \) defined by \( b(t) := \) the barycenter of \( E_i(t) \). Then we want to show that \( b = b(t) \) has also speed \( \leq 1 \).

In order to show this, we need a theorem to deal with the distance between \( b(t) \) and \( b(t') \). From [25], we have the following theorem. This theorem gives an upper bound of the distance between two barycenters by integrating a coupling. Given two probability measures \( \nu_1, \nu_2 \in \mathcal{P}(X) \), we call \( \tilde{\nu} \in \mathcal{P}(X^2) \) a coupling of \( \nu_1 \) and \( \nu_2 \) if \( \tilde{\nu}(A \times X) = \nu_1(A) \) and \( \tilde{\nu}(X \times A) = \nu_2(A) \) for \( \forall A \in \mathcal{B}(X) \).

**Theorem 42.** [25, Th. 6.3] Let \( X \) be a CAT(0) space. If two probability measures \( \nu_1 \) and \( \nu_2 \) satisfy \( \int_X d(x_0, y)\nu_1(dy) < \infty \) for some \( x_0 \in X \) and \( \tilde{\nu} \) is a coupling of \( \nu_1 \) and \( \nu_2 \), then

\[
d(b(\nu_1), b(\nu_2)) \leq \int_{X^2} d(x, y)\tilde{\nu}(dxdy).
\]

**Proposition 43.** Let \( X \) be a CAT(0) space. Given \( n \) evaders \( E_i = E_i(t) \) with speed \( \leq 1 \) in \( X \), then the barycenter curve \( b = b(t) \) is 1-Lipschitz.

**Proof.** Let \( \tilde{\nu} = \frac{1}{n} \sum_{i=1}^{n} \delta_{E_i(t), E_i(t')} \). Then \( \tilde{\nu} \) is the coupling of \( \nu_1 \) and \( \nu_1' \), where \( \nu_1 = \frac{1}{n} \sum_{i=1}^{n} \delta_{E_i(t)} \) and \( \nu_1' = \frac{1}{n} \sum_{i=1}^{n} \delta_{E_i(t')} \).

Obviously, \( \nu_1 \) satisfies \( \int_X d(x_0, y)\nu_1(dy) < \infty \) for some \( x_0 \in X \). Since \( b(t) = b(\nu_1) \) and \( b(t') = b(\nu_1') \) from Lemma 41, by applying \( \nu_1, \nu_1' \), and \( \tilde{\nu} \) to Theorem 42, we have

\[
d(b(t), b(t')) \leq \frac{1}{n} \sum_{i=1}^{n} d(E_i(t), E_i(t')).
\]

Since each evader \( E_i \) has speed \( \leq 1 \), \( d(E_i(t), E_i(t')) \leq |t - t'| \). Then we have

\[
d(b(t), b(t')) \leq |t - t'|.
\]

\[\square\]

This proof shows that the barycenter curve of curves with speed \( \leq 1 \) has also speed \( \leq 1 \). By letting \( Y \) of Corollary 38 be a barycenter curve \( b \), we obtain

**Theorem 44.** Let \( X \) be a CAT(0) space. Given \( n \) evaders \( E_i = E_i(t) \) with speed \( \leq 1 \) in \( X \), there is a unique continuous pursuit curve \( P = P(t) \) chasing the barycenter curve \( b \) of evaders.
Proof. Since \( b \) has speed \( \leq 1 \) by Proposition 43, we take the barycenter curve \( b \) as \( Y \) of Corollary 38.

\[ \square \]

7. Lytchak’s gradient curves

In this section we consider time independent \( F \) and two definitions of gradient curves. One is Lytchak’s definition (see [19]) and the other is the definition we have been using. We first introduce Lytchak’s definition which applies in a very general setting. We then see that Mayer gets Lytchak’s gradient curves. We then show that in our setting Lytchak’s gradient curves are in fact gradient curves in the sense we have used them.

Let us start by defining an absolute gradient of \( F \) at \( x \) from [20] and [24] for the downward case.

**Definition 45.** Let \( X \) be a \( \text{CAT} (0) \) space. For a function \( F : X \to \mathbb{R} \) and \( x \in X \), define the absolute gradient \( |\nabla_{-}F|(x) \) of \( F \) at \( x \) by

\[
|\nabla_{-}F|(x) := \max \left\{ \limsup_{y \to x} \frac{F(x) - F(y)}{d(x,y)}, 0 \right\}.
\]

The following condition is sufficient for the set \( \{ x \in X : |\nabla_{-}F|(x) \neq 0 \} \) of non-critical points to be open:

**Definition 46.** [19], [24] Let \( X \) be a \( \text{CAT} (0) \) space. For a function \( F : X \to \mathbb{R} \), \( F \) has semi-continuous absolute gradients if

\[
\liminf_{y \to x} |\nabla_{-}F|(y) \geq |\nabla_{-}F|(x)
\]

for all \( x \in X \).

By Definition 45, we know:

**Lemma 47.** If \( F \) is locally Lipschitz and \( \lambda \)-convex on a \( \text{CAT} (0) \) space \( X \), then

\[
|||\nabla_{x}(\nabla_{-}F)||| = |\nabla_{-}F|(x)
\]

for any \( x \in X \).

Now, we can give the definition of gradient curves on metric spaces.

**Definition 48.** [19] Let \( X \) be a \( \text{CAT} (0) \) space. For a function \( F : X \to \mathbb{R} \) having semi-continuous absolute gradients, a curve \( m : [0,a) \to X \) is called the (time-independent) gradient curve of \( F \) if for all \( t \in [0,a) \),

\[
\lim_{\epsilon \to 0^+} \frac{d(m(t + \epsilon), m(t))}{\epsilon} = |\nabla_{-}F|(m(t))
\]

and

\[
\lim_{\epsilon \to 0^+} \frac{F \circ m(t + \epsilon) - F \circ m(t)}{\epsilon} = -(|\nabla_{-}F|(m(t)))^2.
\]

Next we look at Mayer’s Theorem from [20]:

**Theorem 49.** [20, Th. 1.13, Th. 2.17, Cor 2.18 and Prop. 2.25] Let \( X \) be a \( \text{CAT} (0) \) space. For \( x_0 \in X \) and a function \( G : X \to \mathbb{R} \), assume that

1) \( G \) is lower semicontinuous,

2) \( G \) is \( \lambda \)-convex.

For any \( y \in X \), let

\[
\begin{align*}
A &= - \min(0, \liminf_{d(x,y) \to \infty} \frac{G(x)}{d^2(x,y)}), \\
I_A &= \begin{cases} 
(0, \infty) & \text{if } A = 0, \\
(0, \frac{1}{16A}) & \text{if } A > 0.
\end{cases}
\end{align*}
\]
Then there is a unique (time-dependent) gradient curve \( m : I_A \to X \) of \( G \) as in Definition 48 such that \( \lim_{t \to 0} m(t) = x_0 \) and \( G(m(t)) \leq G(x_0) \).

Note that \( A \) is independent of \( y \) because of the triangle inequality.

In Theorem 49, it may not happen that the curve \( m \) has right-side tangent vectors \( m'(t+) \) for all \( t \in I_A \). But if \( G \) is not only lower semicontinuous but in fact locally Lipschitz, we show below that the curve \( m : I_A \to X \) which we get by Theorem 49 is the gradient curve such that there exists its right-side tangent vector at every time and it must be equal to a gradient vector.

**Proposition 50.** Let \( X \) be a \( \text{CAT}(0) \) space and \( G : X \to \mathbb{R} \) be locally Lipschitz and \( \lambda \)-convex. Then for the curve \( m : I_A \to X \) which we get by Theorem 49, there exists a right-side tangent vector \( m'(t+) \) at \( t \) and it is equal to \( \nabla m(t)(-G) \) for all \( t \in I_A \).

**Proof.** For \( x := m(t) \), we have the unique downward gradient \( \nabla_x(-G) \in T_x \) from Lemma 12. Then let \( v \) be the gradient \( \nabla_x(-G) \) and \( w_i \) be the tangent in \( T_x \) of a geodesic \([m(t)m(t_i)]\) for any \( t_i > t \). Then it follows from (7.1) and (7.2) that as \( i \to \infty \),

\[
\lim_{i \to \infty} ||w_i|| \to ||\nabla_x G(x)|| \quad \text{and} \quad d_x G(w_i) \to -||\nabla_x G(x)||^2.
\]

If \( v = o_x \), our proof is done since \( ||w_i|| \) goes to 0. Otherwise, by Definition 11, we get

\[
d_x G(w_i) \geq -\langle w_i, v \rangle = -||w_i|| ||v|| \cos \theta_i \quad \text{and} \quad d_x G(v) = -||v||^2
\]

where \( \theta_i \) be the angle between \( w_i \) and \( v \). Then we obtain

\[
-\frac{d_x G(w_i)}{||w_i||} \leq \cos \theta_i.
\]

Since \( ||v|| = ||\nabla_x G(x)|| \) by Definition 45, the left side becomes 1 by (7.4). Then \( \theta_i \) goes to zero as \( i \to \infty \). Our proof is finished.

**ACKNOWLEDGEMENTS**

The author would like to express his gratitude to Prof. S. Alexander and Prof. R. Bishop for their suggestions and advice. He appreciates valuable comments from Prof. C. Croke, Prof. V. Kapovich, Prof. A. Lytchak and D. Lipsky. He is grateful to Prof. R. Ghrist for his support during the preparation of this article. The material in this paper is part of the author’s thesis [13]. He gratefully acknowledge support from the ONR Antidote MURI project, grant no. N00014-09-1-1031.

**REFERENCES**

[1] S. Alexander, D. Berg and R. Bishop, Geometric curvature bounds in Riemannian manifolds with boundary, Trans. Amer. Math. Soc., 339, (1993), 703-716.
[2] S. Alexander, R. Bishop and R. Christ, Total curvature and simple pursuit on domains of curvature bounded above, Geom. Dedicata, 149, (2010), 275-290.
[3] S. Alexander, R. Bishop and R. Ghrist, Pursuit and evasion in non-convex domains of arbitrary dimension, Proc. Robotics Systems and Science, (2006).
[4] S. Alexander, V. Kapovitch and A. Petrunin, Alexandrov meets Kirszbraun, Proc. Gokova Geometry-Topology Conference 2010, S. Akbulut, D. Auroux, T. Onder, eds., International Press, (2011), 88-109.
[5] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the Wasserstein space of probability measures, Birkhäuser, 2004.
[6] M. Bramson, K. Burdzy and W. Kendall, Shy couplings, CAT(0) spaces, and the lion and man, Ann. Probab., 41, (2013), 744-784.
[7] M. Bridson and A. Haefliger, Metric Spaces of Non-positive Curvature, Springer-Verlag, 1999.
[8] D. Burago, Y. Burago and S. Ivanov, A Course in Metric Geometry, Graduate Studies in Mathematics, Vol. 33, Amer. Math. Soc., Providence, RI, 2001.

[9] P. Clément, An Introduction to Gradient Flows in Metric Spaces, Report of the Mathematical Institute Leiden, MI-2009-09, University of Leiden, 2009.

[10] M. Gromov, Hyperbolic groups, Essays in group theory (S. M. Gersten, ed), Springer Verlag, MSRI Publ., 8, (1987), 75-263.

[11] N. Hovakimyan and A. Melikyan, Geometry of pursuit-evasion on second order rotation surfaces, Dynamics and Control, 10(3), (2000), 297-312.

[12] J. Jost, Equilibrium maps between metric spaces, Calc. Var., 2, (1994), 173-204.

[13] C. Jun, Pursuit-evasion and time-dependent gradient flow in singular spaces, PhD thesis, University of Illinois at Urbana-Champaign, (2012).

[14] C. Jun, Continuous pursuit curves on $\text{CAT}(K)$ spaces, Geom. Dedicata, to appear.

[15] H. Kim and N. Masmoudi, Generating and Adding Flows on Locally Complete Metric Spaces, J. Dyn. Diff. Equat., 25, (2013), 231-256.

[16] B. Kleiner, The local structure of length spaces with curvature bounded above, Math. Z., 231, (1999), 409-456.

[17] A. Kovshov, The simple pursuit by a few objects on the multidimensional sphere, Game Theory and Applications II, L. Petrosjan and V. Mazalov, eds., Nova Science Publ., (1996), 27-36.

[18] U. Lang and V. Schroeder, Kirszbraun’s theorem and metric spaces of bounded curvature, Geom. Funct. Anal., 7, (1997), 535-560.

[19] A. Lytchak, Open map theorem for metric spaces, St. Petersburg Math. J., 16, (2005), 1017-1041.

[20] U. F. Mayer, Gradient flows on nonpositively curved metric spaces and harmonic maps, Comm. Anal. Geom., 6, (1998), 199-253.

[21] A. Melikyan, Generalized Characteristics of First Order PDEs, Birkhauser, 1998.

[22] I. G. Nikolaev, The tangent cone of an Alexandrov Space of curvature $\leq K$, Manuscripta Math., 86, (1995), 683-689.

[23] A. Petrunin, Semiconcave functions in Alexandrov’s geometry, Surveys in Differential Geometry XI, (2007), 137-201.

[24] C. Plaut, Metric spaces of curvature $\geq K$, in Handbook of geometric topology, North-Holland, (2002), 819-898.

[25] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002), 357-390. Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003.