Metric and Topological Properties of Paths and Graphs Under the Fréchet Distance

Erin Chambers [*], Brittany Terese Fasy [†], Benjamin Holmgren [†], Sushovan Majhi [‡], Carola Wenk [§]

May 2022

Warning: This manuscript is a preprint and has not yet been published! We are not quite sure where to send it, and it is still being improved upon... If you have any comments or find any typos, please send them to benjamin.holmgren1@student.montana.edu

Abstract

The Fréchet distance is often used to measure distances between paths, with applications arising in road network and GPS trajectory analysis. The Fréchet distance can also be used to study distances between two copies of the same graph embedded or immersed in a metric space. In this paper, we examine topological properties of spaces of paths and of graphs mapped to $\mathbb{R}^n$ under the Fréchet distance. In particular, we prove whether or not these spaces and metric balls in these spaces are path-connected.

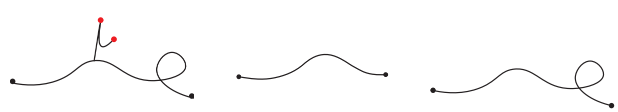
1 Introduction

One-dimensional data in a Euclidean ambient space is heavily studied in the computational geometry and machine learning literature. Perhaps most notably, this form of data is central to applications in GPS trajectory and road network analysis [2, 10, 12, 23]. We attempt to build a theoretical foundation for these application areas by investigating spaces of paths and graphs in $\mathbb{R}^n$ and their metric and topological properties. To compare elements within these types of stratified spaces, the Fréchet distance is arguably the preeminent distance measure, and accounts for the connectivity of the paths or graphs being compared [20, 18, 7, 8, 9, 3, 6, 1, 15, 4, 14, 22, 11, 16, 10, 13, 12]. Moreover, in recent years the Fréchet distance has been demonstrated as a practical comparison measure between such spaces of paths [6, 13], and between graphs [16, 18, 9]

[*]Saint Louis University
[†]Montana State University
[‡]University of California, Berkeley
[§]Tulane University

in both the discrete and continuous context. The motivation for such work is simple: when learning these types of network representations, for example in map matching [10, 12] or industrial GIS settings [23], we want to know the theoretical limitations of an approach. Additionally, nearest neighbor methods using the Fréchet distance are well studied [21, 19], and models have even been introduced to learn the Fréchet distance between trajectories [5]. Our work provides insight toward the theoretical limitations of such approaches. This is to say, if one wants to compare multiple learned outputs from different models, can every output *possibly* be an accurate topological representation of the desired underlying network? Is the Fréchet distance necessarily feasible to use on spaces of paths or graphs in $\mathbb{R}^n$, or could optimization methods conducted on these networks fall apart under perturbation?

In this paper, we define the Fréchet distance among paths and graphs in the most general context possible, so that these results may be applied to the fullest extent within different variants and applications of the Fréchet distance to one dimensional data. Using open balls under the Fréchet distance to generate a topology on sets of one dimensional data mapped in $\mathbb{R}^n$, we study the metric and topological properties of the induced spaces. In particular, we work with three classes of paths: the set $\Pi_{\mathcal{C}}$ of all paths in $\mathbb{R}^n$, the set $\Pi_{\mathcal{E}}$ of all paths in $\mathbb{R}^n$ that are embeddings (i.e., mappings that are homeomorphisms onto the image), and the set $\Pi_{\mathcal{I}}$ of all paths in $\mathbb{R}^n$ that are immersions (local embeddings). See Figure 1 for examples of paths in $\mathbb{R}^2$. In addition, we study three analogous spaces of graphs: the sets $\mathcal{G}_{\mathcal{C}}$, $\mathcal{G}_{\mathcal{I}}$, and $\mathcal{G}_{\mathcal{E}}$ of continuous mappings, immersions, and embeddings respectively. This paper establishes the core mathematical properties of the Fréchet distance and the topological spaces that it induces on graphs and paths in Euclidean space.



(a) Continuous (b) Embedding (c) Immersion

Figure 1: Example of paths continuously mapped, embedded, and immersed in $\mathbb{R}^2$. The space of continuous mappings allows arbitrary self-intersection on a path including backtracking (which occurs at the two red points); embeddings must induce homeomorphisms onto their image; and immersions are locally embeddings.

2 Background

In this section, we establish the definitions and notation from geometry and topology used throughout. For a review of fundamental definitions in computational topology, we refer readers to Edelsbrunner and Harer [17]. Let $\mathbb{X}$ and $\mathbb{Y}$ be topological spaces.

Definition 1 (Types of Continuous Maps)**.** *A map $\alpha\colon \mathbb{X} \to \mathbb{Y}$ is called* continuous *if for each open set $U \subset \mathbb{Y}$, $\alpha^{-1}(U)$ is open in $\mathbb{X}$. We call α an* embedding *if α is injective. Alternatively, an* embedding *is a continuous map that is a homeomorphism onto its image. If α is locally an embedding (that is, for any $x \in X$ there exists an open neighborhood N containing x such that $\alpha|_N$ is injective), then we say that α is an* immersion*.*

Given a set $\mathbb{X}$ and a notion of distance between objects in $\mathbb{X}$, we can topologize $\mathbb{X}$.

Definition 2 (The Open Ball Topology)**.** *Let $\mathbb{X}$ be a set and $d\colon \mathbb{X}\times\mathbb{X} \to \mathbb{R}_{\geq 0}$ a distance function. For each $r > 0$ and $x \in \mathbb{X}$, let $\mathbb{B}_d(x, r) := \{y \in \mathbb{X} \mid d(x, y) < r\}$. The open ball topology on $\mathbb{X}$ with respect to d is the topology generated by $\{\mathbb{B}_d(x, r) \mid x \in \mathbb{X}, r > 0\}$.*

In words, $\mathbb{B}_d(x, r)$ denotes the open ball of radius r centered at x with respect to the metric d. We use these open balls to generate a topology on $\mathbb{X}$, allowing x to range over $\mathbb{X}$ and r to range over all positive real numbers.

Definition 3 (Path)**.** *A* path *in a topological space $\mathbb{X}$ between two elements $a, b \in \mathbb{X}$, is defined to be a continuous map $\Gamma : [0, 1] \to \mathbb{X}$ where $\Gamma(0) = a$, and $\Gamma(1) = b$.*

With the definition of paths, we define the main property of interest in this paper: path-connectivity.

Definition 4 (Path-Connectivity)**.** *A topological space $\mathbb{X}$ is called* path-connected *if there exists a* path *between any two elements in $\mathbb{X}$.*

And, *length* of a path in a general metric space is given by:

Definition 5 (Length)**.** *Let $(\mathbb{X}, d)$ be a metric space and let γ be a* path *in $(\mathbb{X}, d)$. Let $0 = t_0 < t_1 < ... < t_n = 1$ and construct a partition $p = \{t_0, t_1, ...t_n\}$ of $[0, 1]$. The* length *$L_d(\gamma)$ of γ is defined as the supremum over P, the set of all such partitions p:*

$$L_d(\gamma) := \sup_P \sum_{i=1}^{n} d(\gamma(t_i), \gamma(x_{i-1})).$$

Finally, many paths we encounter have finite lengths. We call these paths *rectifiable.* Moreover, we call a one-dimensional topological space (i.e., a graph) $\mathbb{X}$ rectifiable if there exists a cover of $\mathbb{X}$ such that every element in the open cover is a rectifiable path.

Sets of Paths in $\mathbb{R}^n$ Throughout this paper, we focus on paths and graphs in $\mathbb{R}^n$. We denote the set of all rectifiable paths in $\mathbb{R}^n$ by $\Pi_{\mathcal{C}}$, the set of all rectifiable, immersed paths in $\mathbb{R}^n$ by $\Pi_{\mathcal{I}}$, and the set of all rectifiable, embedded paths in $\mathbb{R}^n$ by $\Pi_{\mathcal{E}}$. Note that $\Pi_{\mathcal{E}} \subsetneq \Pi_{\mathcal{I}} \subsetneq \Pi_{\mathcal{C}}$ by definition. For examples, we again refer to Figure 1.

We now define the standard Fréchet distance between paths in $\mathbb{R}^n$, as defined in Alt and Godau [3].

Definition 6 (The Fréchet Distance for Paths)**.** *The* Fréchet distance $d_{FP}\colon \Pi_{\mathcal{C}} \times \Pi_{\mathcal{C}} \to \mathbb{R}_{\geq 0}$ *between* $\gamma_1, \gamma_2 \in \Pi_{\mathcal{C}}$ *is defined as:*

$$d_{FP}(\gamma_1, \gamma_2) := \inf_{r\colon [0,1]\to[0,1]} \max_{t\in[0,1]} ||\gamma_1(t) - \gamma_2(r(t))||_2,$$

where r ranges over all reparameterizations of the unit interval (that is, homeomorphisms such that $r(0) = 0$ and $r(1) = 1$), and $||\cdot||_2$ denotes the Euclidean norm.

Sets of Graphs in $\mathbb{R}^n$ We define the analogous spaces of (multi-)graphs. In what follows, we define a *graph* G as a finite set of vertices V and a finite set of edges E, and denote it as $G = (V, E)$. Self-loops and multiple edges between a pair of vertices are allowed in our setting. We topologize a graph by thinking of it as a CW complex; that is, each open edge is homeomorphic to the open interval $(0, 1) \subset \mathbb{R}$ with endpoints glued to vertices. As with paths, we have three sets of interest. For a graph G, these sets are:

1. $\mathcal{G}_{\mathcal{C}}(G)$ is the set of all continuous, rectifiable maps $\phi\colon G \to \mathbb{R}^n$.
2. $\mathcal{G}_{\mathcal{I}}(G)$ is the set of rectifiable immersions of G into $\mathbb{R}^n$.
3. $\mathcal{G}_{\mathcal{E}}(G)$ is the set of rectifiable embeddings of G into $\mathbb{R}^n$.

Taking the set of all graphs As with paths, by design $\mathcal{G}_{\mathcal{E}} \subsetneq \mathcal{G}_{\mathcal{I}} \subsetneq \mathcal{G}_{\mathcal{C}}$.

With these spaces in hand, we extend the Fréchet distance for paths by providing a Fréchet distance for graphs continuously mapped into $\mathbb{R}^n$:

Definition 7 (Graph Fréchet Distance)**.** *Let G and H be graphs, and let $\phi\colon G \to \mathbb{R}^n$ and $\psi\colon H \to \mathbb{R}^n$ be continuous, rectifiable mappings of graphs. Given any homeomorphism $h\colon G \to H$, we say that the* induced L_∞ distance *between the maps ϕ and $\psi \circ h$ is $||\phi - \psi \circ h||_\infty = \max_{x\in G} |\phi(x) - \psi(h(x))|$. With this distance in hand, we define the Fréchet distance between (G, ϕ) and (H, ψ) by minimizing over all homeomorphisms:*[1]

$$d_{FG}\left((G,\phi),(G,\psi)\right) := \begin{cases} \inf_h ||\phi - \psi \circ h||_\infty & G \cong H. \\ \infty & \textit{otherwise.} \end{cases}$$

For simplicity of exposition, when $G = H$ and is understood by context, we often write the LHS of this equation as $d_{FG}(\phi, \psi)$.

[1]Other generalizations of the Fréchet distance minimize over all "orientation-preserving" homeomorphisms, which can be defined in several ways for stratified spaces. We drop this requirement in our definition.

Observation 8. *If $G = [0,1]$, then the relationship between the Fréchet distance between two paths $\alpha, \beta \colon [0,1] \to \mathbb{R}^n$ is as follows:*

$$d_{FG}\left(\alpha, \beta\right) = \min\left\{d_{FP}(\alpha, \beta), d_{FP}(\alpha, \beta^{-1})\right\},$$

where $\beta^{-1} \colon I \to \mathbb{R}^n$ is defined by $\beta^{-1}(t) = \beta(1-t)$.

Note also that elements of $\mathcal{G}_{\mathcal{C}}, \mathcal{G}_{\mathcal{I}}$, and $\mathcal{G}_{\mathcal{E}}$ are equivalent (with graph Fréchet distance zero) if there exists a homeomorphism between the graphs that serves as a way to “reparameterize" the maps.

Contextualizing Spaces of Paths and Graphs and Their Properties The spaces of paths and graphs defined are relevant in machine learnixng domains where a combinatorial graph is mapped in $\mathbb{R}^n$, and the positioning of both vertices and edges may matter. Most notably, applications in GPS trajectory analysis generally output graph immersions or embeddings (depending on the consideration of bridges or tunnels) in either two or three dimensions. These objects fit directly as elements within our definitions of the sets $\Pi_{\mathcal{I}}, \Pi_{\mathcal{E}}$ and $\mathcal{G}_{\mathcal{I}}, \mathcal{G}_{\mathcal{E}}$ [2, 10, 12, 23]. Alternatively, many graph "embedding" methods (embedding here not referring to the mathematical term) have been introduced throughout machine learning for dimensionality reduction, where the idea is to map a graph with m vertices (i.e. an mxm adjacency matrix) into a lower dimensional space, where graph-based distances between vertices are preserved by Euclidean distance. Here, the spaces $\Pi_{\mathcal{C}}$ and $\mathcal{G}_{\mathcal{C}}$ apply by definition, since the images of edges are not addressed in these definitions [25, 24]. However, for a meaningful projection (where the images of every vertex are unique), there is a canonical graph immersion (given in Appendix A.1), demonstrating $\Pi_{\mathcal{I}}$ and $\mathcal{G}_{\mathcal{I}}$ also apply in this setting. Additionally, in dimensionality reduction use cases, it is reasonable to expect higher dimensional Euclidean ambient spaces than $\mathbb{R}^3$.

Basic path-connectedness properties are chosen as the subject of study for a number of reasons. As these spaces arise often in application areas throughout machine learning and computational geometry, it is reasonable to inquire about their fundamental behavior. When nearest neighbor methods using the Fréchet distance [21, 19] are applied to a road network, for example, will they necessarily produce outputs topologically consistent to the original network? Will transforming one network into another necessarily increase the Fréchet distance along the way? Do Fréchet-like distances come with unforeseen complications when comparing graph mappings widely used for dimensionality reduction [25, 24]? Is the Fréchet distance necessarily meaningful to approximate using machine learning, or is it possible that unforeseen discontinuities could arise in this formulation, particularly when extended to graphs [5]? Our paper attempts to address these fundamental questions by proof of basic metric and topological properties.

3 Metric Properties

We now address perhaps the first natural question when studying a distance acting upon a space: Is this distance a metric? We prove that d_{FG} forms a pseudo-metric acting on $\mathcal{G}_{\mathcal{C}}, \mathcal{G}_{\mathcal{I}}$, and $\mathcal{G}_{\mathcal{E}}$.

It is well known that the Fréchet distance for paths is a pseudo-metric [20, 3], (that is, satisfying all metric properties except for separability, which is mended easily by defining an equivalence class among paths with Fréchet distance zero) and we provide proof of this property for the Fréchet distance among graphs. The following proof is naturally restricted to d_{FP} acting on $\Pi_{\mathcal{C}}, \Pi_{\mathcal{I}}$, and $\Pi_{\mathcal{E}}$.

Theorem 9 (Metric Properties of $(\mathcal{G}_{\mathcal{C}}(G), d_{FG})$)**.** *If G is a graph, d_{FG} is a pseudo-metric on $\mathcal{G}_{\mathcal{C}}(G)$.*

Proof. We demonstrate that the space $\mathcal{G}_{\mathcal{C}}(G)$ under the Fréchet distance is a pseudo-metric:

First, we prove identity (that $(d_{FG}(\mathbf{G}, \mathbf{G}) = 0)$). Let $\mathbf{G} = (G, \phi) \in \mathcal{G}_{\mathcal{C}}(G)$. Then

$$d_{FG}(\mathbf{G}, \mathbf{G}) = \inf_{h} \max_{x \in G} |\phi(x) - \phi(h(x))| \leq \inf_{h} \max_{x \in G} |\phi(x) - \phi(id(x))| = 0.$$

Next, we prove symmetry (that $(d_{FG}(\mathbf{G}_1, \mathbf{G}_2) = d_{FG}(\mathbf{G}_2, \mathbf{G}_1))$. Let $\mathbf{G}_1 = (G, \phi), \mathbf{G}_2 = (G, \psi) \in \mathcal{G}_{\mathcal{C}}(G)$. Examine $d_{FG}(\mathbf{G}_1, \mathbf{G}_2) = \inf_{h} ||\phi - \psi \circ h||_\infty$. This could just as well be written $\inf_{h^{-1}} ||\phi \circ h^{-1} - \psi||_\infty$, since h^{-1} is by definition a homeomorphism, which is precisely $d_{FG}(\mathbf{G}_2, \mathbf{G}_1)$.

Next, we examine the failure of seperability (that $(d_{FG}(\mathbf{G}_1, \mathbf{G}_2) = 0 \not\Longrightarrow \mathbf{G}_1 = \mathbf{G}_2)$). Let $\mathbf{G}_1 = (G, \phi), \mathbf{G}_2 = (G, \psi) \in \mathcal{G}_{\mathcal{C}}(G)$. Suppose $d_{FG}(\mathbf{G}_1, \mathbf{G}_2) = 0$. Then there exists a homeomorphism $h : G \to G$ satisfying $||\phi - \psi \circ h||_\infty = 0$. Then this requires $\phi = \psi \circ h$, which does not require that $\phi = \psi$ and therefore it may be the case that $(G, \phi) \neq (G, \psi)$. So $\mathbf{G}_1 \neq \mathbf{G}_2$ necessarily, and separability is not fulfilled. However, if we define the canonical equivalence class in the space with $\mathbf{G}_1 \sim \mathbf{G}_0 \iff d_{FG}(\mathbf{G}_1, \mathbf{G}_0) = 0$, then we maintain separability and induce a metric.

Finally, we prove subadditivity (the triangle inequality). Let $\mathbf{G}_1 = (G, \phi_1), \mathbf{G}_2 = (G, \phi_2)$, and $\mathbf{G}_3 = (G, \phi_3)$ be in $\mathcal{G}_{\mathcal{C}}(G)$. Fix an $\epsilon > 0$. Then, there exists reparameterization h such that:

$$||\phi_1 \circ h - \phi_2||_\infty \leq d_{FG}(G_1, G_2) + \frac{\varepsilon}{2}.$$

Similarly, then there exists reparameterization i such that:

$$||\phi_2 - \phi_3 \circ i||_\infty \leq d_{FG}(G_2, G_3) + \frac{\varepsilon}{2}.$$

Now,

$$\begin{aligned}
d_{FG}(\mathbf{G}_1, \mathbf{G}_3) &= \inf_{f,g} ||\phi_1 \circ f - \phi_3 \circ g||_\infty \\
&\leq ||\phi_1 \circ h - \phi_3 \circ i||_\infty \\
&= ||\phi_1 \circ h - \phi_2 + \phi_2 - \phi_3 \circ i||_\infty \\
&\leq ||\phi_1 \circ h - \phi_2||_\infty + ||\phi_2 - \phi_3 \circ i||_\infty, \text{ by the triangle inequality of the inf norm.} \\
&\leq d_{FG}(G_1, G_2) + \frac{\varepsilon}{2} + d_{FG}(G_2, G_3) + \frac{\varepsilon}{2} \\
&= d_{FG}(G_1, G_2) + d_{FG}(G_2, G_3) + \varepsilon.
\end{aligned}$$

Since ε is chosen arbitrarily, we conclude that $d_{FG}(\mathbf{G}_1, \mathbf{G}_3) \leq d_{FG}(G_1, G_2) + d_{FG}(G_2, G_3)$.

Hence, the space $\left(\mathcal{G}_{\mathcal{C}}(G), d_{FG}\right)/_{\sim}$ maintains the properties of a metric space.

□

Corollary 10 (Metric Extension to $\mathcal{G}_{\mathcal{I}}, \mathcal{G}_{\mathcal{E}}$)**.** *The Fréchet distance between graphs is an extended metric on $\mathcal{G}_{\mathcal{I}}$ and $\mathcal{G}_{\mathcal{E}}$ modulo $\sim$, the equivalence relation accompanying graphs with Fréchet distance zero. If G is a graph, then the Fréchet distance between graphs is a metric on the quotient spaces $\mathcal{G}_{\mathcal{I}}(G)/_{\sim}$ and on $\mathcal{G}_{\mathcal{E}}(G)/_{\sim}$.*

Proof. Recalling that $\mathcal{G}_{\mathcal{C}} \subsetneq \mathcal{G}_{\mathcal{I}} \subsetneq \mathcal{G}_{\mathcal{E}}$, it follows from the metric properties of $\left(\mathcal{G}_{\mathcal{C}}(G), d_{FG}\right)/_{\sim}$ that the subspaces $\left(\mathcal{G}_{\mathcal{I}}(G), d_{FG}\right)/_{\sim}$ and $\left(\mathcal{G}_{\mathcal{E}}(G), d_{FG}\right)/_{\sim}$ each form a metric space. □

Corollary 11 (Metric Properties of d_{FP})**.** *The Fréchet distance between paths is an extended pseudo-metric on $\Pi_{\mathcal{C}}, \Pi_{\mathcal{I}}$ and $\Pi_{\mathcal{E}}$.*

Proof. The proof of Theorem 9 is naturally restricted to paths, taking $G = [0, 1]$. □

4 Path-Connectedness Property

We now examine the path-connectedness properties in an attempt to make basic theoretical guarantees about the topological characteristics of the spaces $\Pi_{\mathcal{C}}, \Pi_{\mathcal{I}}, \Pi_{\mathcal{E}}$ and $\mathcal{G}_{\mathcal{C}}, \mathcal{G}_{\mathcal{I}}, \Pi_{\mathcal{E}}$.

4.1 Continuous Mappings

We start with the broadest spaces of paths and graphs: those which are continuously mapped into $\mathbb{R}^n$.

Theorem 12 (Continuous Mappings of Paths)**.** *The topological space $(\Pi_{\mathcal{C}}, d_{FP})$ is path-connected.*

Proof. Let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{C}}$. Let $\Gamma : [0,1] \to \Pi_{\mathcal{C}}$ be the linear interpolation from $\Gamma(0) = \gamma_0$ to $\Gamma(1) = \gamma_1$ along the pairwise matchings defining $d_F(\gamma_0, \gamma_1)$. We leave rigorous definition of linear interpolation to Definition 31, stemming from the linear combinations of graphs in Definition 30. Since $\Gamma(t) \in \Pi_{\mathcal{C}}$ for each t, Γ is well-defined. Additionally, due to Lemma 32 taking $G = [0,1]$, Γ is continuous. $\square$

Corollary 13 (Metric Balls in $(\Pi_{\mathcal{C}}, d_{FP})$)**.** *Metric balls in the space* $(\Pi_{\mathcal{C}}, d_{FP})$ *are path-connected.*

Proof. By design, direct interpolation never increases the Fréchet distance. Let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{C}}$, and let $\delta := d_{FP}(\gamma_0, \gamma_1)$. Let Γ be the linear interpolation from γ_0 and γ_1 in $\Pi_{\mathcal{C}}$. As t increases, the pairwise matchings defining $d_{FP}(\Gamma(t), \gamma_1)$ decrease in length by construction and so, for all $t \in (0,1)$, we have $d_{FP}(\Gamma(t), \gamma_1) \leq \delta$, which means that $\Gamma(t) \in \mathbb{B}_{d_{FP}}(\gamma_1, \delta)$. Hence, metric balls in the space $(\Pi_{\mathcal{C}}, d_{FP})$ are path-connected. $\square$

Theorem 14 (Continuous Mappings of Graphs)**.** *Let G be a graph. The metric space $(\mathcal{G}_{\mathcal{C}}(G), d_{FG})$ is path-connected. Moreover, the connected components of the extended metric space $(\mathcal{G}_{\mathcal{C}}, d_{FG})$ are in one-to-one correspondence with the homeomorphism classes of graphs.*

Proof. Let $\mathbf{G}_0, \mathbf{G}_1 \in \mathcal{G}_{\mathcal{C}}(G)$. Let Γ be the linear interpolation from $\mathbf{G}_0$ to $\mathbf{G}_1$, defined by interpolating along the pairwise matchings defining $d_{FG}(\mathbf{G}_0, \mathbf{G}_1)$. For readability, we leave rigorous definition to Definition 31 in Appendix B.1). Since Γ is well-defined and continuous, we have a path in $\mathcal{G}_{\mathcal{C}}(G)$ from $\mathbf{G}_0$ to $\mathbf{G}_1$, which means that $(\mathcal{G}_{\mathcal{C}}(G), d_{FG})$ is path-connected. Moreover, suppose $\mathbf{G}_0 = (G, \phi_0)$, $\mathbf{G}_1 = (G', \phi_1) \in \mathcal{G}_{\mathcal{C}}$ for G, G' which are not homeomorphic. Then, $d_{FG}(\mathbf{G}_0, \mathbf{G}_1)$ is undefined, and connected components of the extended metric space $\mathcal{G}_{\mathcal{C}}$ are vacuously in one-to-one correspondence with homeomorphism classes of graphs. $\square$

Corollary 15 (Metric Balls in $(\mathcal{G}_{\mathcal{C}}, d_{FP})$)**.** *Metric balls in the space* $(\mathcal{G}_{\mathcal{C}}(G), d_{FP})$ *are path-connected.*

Proof. Let $\mathbf{G}_0, \mathbf{G}_1 \in \mathcal{G}_{\mathcal{C}}$ and suppose $d_{FG}(\mathbf{G}_0, \mathbf{G}_1) = \delta$. The proof is identical to that in Corollary 13, where it is immediate from observing the path Γ used in Theorem 14 that $\Gamma(t) \in \mathbb{B}_{d_{FG}}(\mathbf{G}_1, \delta)$ for any $t \in (0,1)$. Again, this is due to the decreasing lengths of leashes defining d_{FG} under the interpolation defining Γ. $\square$

4.2 Immersions

In a path immersion, local injectivity is required. Hence, self-intersections are allowed, but pausing or reversing direction is not allowed. See Figure 2 for examples. To show the path-connectivity of spaces of immersions, the proofs in Theorem 12 and Theorem 14 for continuous mappings *almost* suffice, but these added constraints must be addressed. We formally define such notions below:

(a) Forced Backtracking (b) Constant Map

Figure 2: Examples of paths in $\Pi_{\mathcal{C}}$ but not $\Pi_{\mathcal{I}}$. Figure 2a demonstrates a path with necessary backtracking, and hence a path which is not locally injective at the red point. Figure 2b demonstrates a path which (vacuously) must pause and is therefore not locally injective. For a nontrivial example of such a path, consider any parameterization of a path sending a closed interval to a fixed point.

Definition 16 (Pausing)**.** *We say that a path γ* pauses *in an open interval $(a, b) \subset [0, 1]$ If $\gamma(x) = \gamma(y)$ for every $x, y \in (a, b)$. This forces a degeneracy in the space of immersions, and $\gamma \notin \Pi_{\mathcal{I}}$.*

Definition 17 (Backtracking)**.** *We say that a path γ is* backtracking *at a point $x \in [0, 1]$ if there exists $\delta > 0$ such that for every $\epsilon > 0$ with $\delta > \epsilon$, $\gamma|_{(x-\epsilon,x)} \subset \gamma_{(x,x+\epsilon)}$ or $\gamma|_{(x,x+\epsilon)} \subset \gamma|_{(x-\epsilon,x)}$.*

Here, we show the path-connectivity of $(\Pi_{\mathcal{I}}, d_{FP})$ and $(\mathcal{G}_{\mathcal{I}}, d_{FG})$ in dimension greater than one by introducing a reparameterization maneuver to maintain local injectivity during a forced pause.

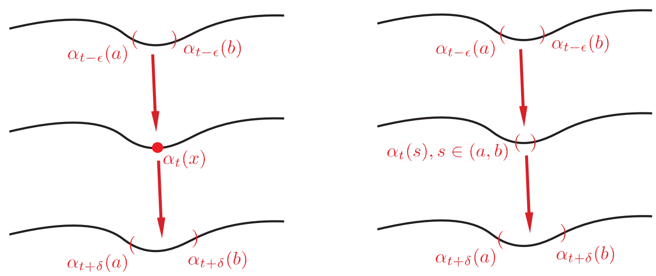
Lemma 18 (Pauses can be Reparameterized)**.** *Suppose at a time $t \in [0, 1]$, a linear interpolation $\Gamma(t)$ in $\Pi_{\mathcal{I}}$ would result in a map γ_t* pausing *on some open interval $(a, b) \subset [0, 1]$. Then there exists a reparameterization of γ_t for sufficiently small $\epsilon > 0$ such that $\Gamma(s) \in \Pi_{\mathcal{I}}$ for all $s \in (t-\epsilon, t+\delta)$, where $\delta > 0$ and $t + \delta$ denotes the time for which a pause concludes in Γ.*

Proof. Denote a pausing point $\gamma_t(x)$ for $x \in (a, b)$, the open interval in the domain of γ_t where the map is not injective. This pause can be subverted by carefully reparameterizing γ_t at time $t - \epsilon$ for sufficiently small $\epsilon > 0$. Namely, we reparameterize each path in $\Pi_{\mathcal{I}}$ given by $\Gamma|_{(t-\epsilon,t+\delta)}$ by inflating an open interval of radius r about the fixed point in γ_t:

$$\gamma_t(s) = \begin{cases} \gamma_t(s * (1 - 2r)) & \text{if } s \in [0, a] \\ \gamma_t(x - (r * s)) & \text{if } s \in (a, a + (b - a)/2] \\ \gamma_t(x + (r * s)) & \text{if } s \in [a + (b - a)/2, b) \\ \gamma_t(s * (1 - 2r) + 2r) & \text{if} s \in [b, 1] \end{cases} \tag{1}$$

Where $r > 0$ is sufficiently small to maintain continuity in $\gamma_t = \Gamma(t)$ for each $t \in (a, b)$. For an example, see Figure 3. □

Observation 19. *Note that such an r must exist, due to the rectifiability constraint of any $\gamma_t \in \Pi_{\mathcal{I}}$.*



(a) Example of a pausing event

(b) Reparameterizing a pause

Figure 3: At a pausing event, we reparameterize a sufficiently small interval of radius r surrounding the degenerate point, thereby maintaining local injectivity and the continuity of all paths in $\Gamma|_{(t-\epsilon,t+\delta)}$.

Direct linear interpolation can also yield degeneracies by creating a singleton in specific circumstances, or by creating a backtracking point. Each of these scenarios require less involved maneuvers, which are sketched in the following theorem, and addressed formally in Lemma 33 and Lemma 34.

Theorem 20 (Path Immersions)**.** *The topological space* $(\Pi_{\mathcal{I}}, d_{FP})$ *of paths immersed in* $\mathbb{R}^n$ *is path-connected iff* $n > 1$.

Proof. That $\Pi_{\mathcal{I}}$ is not path-connected in dimension 1 is trivial from observing that traversing between paths with conflicting orientation must degenerate to a point.

However, if $n > 1$, $\Pi_{\mathcal{I}}$ is path-connected. Let $\Gamma : [0, 1] \to \Pi_{\mathcal{C}}$ be the linear interpolation from $\Gamma(0) = \gamma_0$ to $\Gamma(1) = \gamma_1$, as defined in Definition 31. If $\Gamma[t] \in \Pi_{\mathcal{I}}$ for each $t \in [0, 1]$, we are done. Otherwise, let $T \subset I$ be the set of times that introduce a non-immersion (i.e., $t \in T$ iff $\Gamma(t) \notin \Pi_{\mathcal{I}}$, but $\Gamma(t-\epsilon) \in \Pi_{\mathcal{I}}$ for all ϵ small enough). There are two things that might have happened at t: either an interval collapsed to a point (a pause) or the direction was reversed.

1. Suppose there exists $t \in T$ which corresponds to an interval pausing as in Definition 16, illustrated in Figure 2b. Again, note that a pausing event can occur either if an open interval of $\Gamma(t)$ becomes degenerate, or if $\Gamma(t)$ collapses entirely to a singleton.

   If pausing occurs only on a partial interval of $[0, 1]$, it can be avoided using Lemma 18, where the pausing event is guaranteed to conclude at some time $t + \delta$ for $\delta \geq 0$ since $\gamma_1 \in \Pi_{\mathcal{I}}$, and the interpolation eventually must attain γ_1.

   If a pausing event stems from a full collapse to a singleton (i.e. interpolation occurs between two colinear segments with reverse orientation, and consequently degenerate to a point), the collapse can be circumvented by rotating the segment defined by $\Gamma(t)$, as outlined in Lemma 33.

2. Alternatively, suppose there exists time $t \in T$ which corresponds to backtracking at a point in the path $\Gamma(t)$ according to Definition 17 and Figure 2a. Here, $\Gamma(t)$ can remain in $\Pi_{\mathcal{I}}$ by inflating a ball of radius ϵ for some sufficiently small $\epsilon > 0$ about the backtracking point before it is created. This is outlined formally in Lemma 34, and shown in Figure 7b.

For all $t \in T$, the described moves can be used to subvert potential lapses in injectivity along Γ. Hence, we can construct the desired path Γ by interpolating from α_0 to α_1, and applying the corresponding move at each $t \in T$ to handle pauses or backtracking. Thus, $(\Pi_{\mathcal{I}}, d_{FP})$ is path-connected.

□

Corollary 21 (Metric Balls in $(\Pi_{\mathcal{I}}, d_{FP})$)**.** *Metric balls in $(\Pi_{\mathcal{I}}, d_{FP})$ are path-connected if $n > 1$.*

Proof. Let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{I}}$. As for continuous mappings, paths in the space of immersions as constructed in Theorem 20 need not increase $d_{FP}(\gamma_0, \gamma_1)$. By design, avoiding singleton degeneracies by way of Lemma 33 does not increase the Fréchet distance. Moreover, Lemma 18 and Lemma 34 could potentially increase $d_{FP}(\gamma_t, \gamma_1)$ at some time $t \in [0, 1]$, but in this case the points in question (either a fixed point or a critical backtracking point) can be perturbed slightly in order to no longer define the Fréchet distance. Hence, these moves need not result in $d_{FP}(\gamma_t, \gamma_1) > d_{FP}(\gamma_0, \gamma_1)$, and balls in $\Pi_{\mathcal{I}}$ are path-connected. □

Theorem 22 (Graph Immersions)**.** *The connected components of the extended topological space $(\mathcal{G}_{\mathcal{I}}, d_{FG})$ are in one-to-one correspondence with the homeomorphism classes of graphs.*

Proof. First, we show that for a given graph G, the metric space $(\mathcal{G}_{\mathcal{I}}(G), d_{FG})$ for a graph G is path-connected. We construct Γ identically to Theorem 20, though interpolation and self-crossings occur according to the pointwise matchings among edges, rather than the matchings between individual segments. However, as in Theorem 20 local injectivity can only be violated by pauses, point degeneracies, and backtracking, which can be handled equivalently. □

Corollary 23 (Metric Balls in $(\mathcal{G}_{\mathcal{I}}, d_{FG})$)**.** *Metric balls in the space $(\mathcal{G}_{\mathcal{I}}(G), d_{FG})$ are path-connected.*

Proof. Let $\mathbf{G}_0, \mathbf{G}_1 \in \mathcal{G}_{\mathcal{I}}$ such that $d_{FG}(\mathbf{G}_0, \mathbf{G}_1) = \delta$ and examine the path Γ in Theorem 22. An identical argument to the one in Corollary 21 is sufficient to demonstrate the extension for graphs. That is, that linear interpolation and subsequent handling of possible violations by design never violate that $\Gamma(t) \in \mathbb{B}_{d_{FG}}(\mathbf{G}_1, \delta)$ for every $t \in (0, 1)$, since interpolation and the moves in Lemma 33, Lemma 34, and Lemma 18 never strictly increase the Fréchet distance when extended to graphs. □

4.3 Embeddings

Lastly, we examine the path-connectedness property of the analogous spaces of embeddings.

Theorem 24 (Path Embeddings)**.** *The topological space* $(\Pi_{\mathcal{E}}, d_{FP})$ *is path-connected iff* $n > 1$*.*

In one dimension, $(\Pi_{\mathcal{E}}, d_{FP})$ is not path-connected trivially just as in $(\Pi_{\mathcal{I}}, d_{FP})$.

In higher dimensions, let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{E}}$. There exists a canonical path from γ_0 to γ_1 by condensing each map toward its center until the images are "nearly straight", continuously mapping each image to a straight segment, and then interpolating as in Theorem 12.

Proof. Without loss of generality, we need to construct a continuous $\Gamma : [0,1] \to \Pi_{\mathcal{E}}$ in the extended metric space $(\Pi_{\mathcal{E}}, d_{FG})$ such that $\Gamma(0) = \gamma_0$ and $\Gamma(1) = \gamma_1$. To begin, define $\Gamma_0 : [0,1] \to \Pi_{\mathcal{E}}$, and $\Gamma_1 : [0,1] \to \Pi_{\mathcal{E}}$, by restricting the domains of γ_0, and γ_1, thereby condensing each curve inward:

$$\Gamma_0^s(t) := \gamma_0|_{[s/2, 1-s/2]}(t)$$

$$\Gamma_1^s(t) := \gamma_1|_{[s/2, 1-s/2]}(t)$$

Then, as $t \to 1$, the images of γ_0 and γ_1 encompass an increasingly smaller, and therefore straighter curve in the embedding space. As a consequence of Taylor's theorem, both images must attain some juncture at time t_0^* and t_1^* where $\gamma_0|_{(t_0^*/2, 1-t_0^*/2)}$ and $\gamma_1|_{(t_1^*/2, 1-t_1^*/2)}$ can be continuously straightened in $\Pi_{\mathcal{E}}$ toward the line tangent to the center of each curve. From there, the two straight lines must be made parallel, and then a standard interpolation between straight segments may be used to transform the remaining image of γ_0 to γ_1. Consequently, we obtain the desired Γ by the composition of the condensing maps Γ_0^s and $\Gamma_1^s(t)$, and the straightening and linear interpolation steps once each condensing map has attained the restriction $\gamma_0|_{(t_0^*/2, 1-t_0^*/2)}$ and $\gamma_1|_{(t_1^*/2, 1-t_1^*/2)}$.

Note that the requirement in Section 2 that γ_0 and γ_1 are rectifiable is crucial for the provided construction. Otherwise, there would be no guarantee that one could condense the images of γ_0 and γ_1 to become "straight enough" in order to continuously achieve a straight segment in the space $\Pi_{\mathcal{E}}$. ☐

Theorem 25 (Metric Balls in $(\Pi_{\mathcal{E}}, d_{FP})$)**.** *Metric balls in the space* $(\Pi_{\mathcal{E}}, d_{FP})$ *are path-connected in* $\mathbb{R}^n$ *if* $n \geq 4$*.*

Proof. Let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{E}}$. That metric balls are not connected in one dimension holds vacuously in the space of embeddings. In dimension two and three, metric balls are not necessarily path-connected, due to counterexamples such as those constructed in Lemma 35 and Lemma 36 and illustrated in Figure 4. In dimension four or greater, metric balls are path-connected due to Lemma 37. ☐

Figure 4: Two embedded paths γ_0, γ_1 in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively, for which constructing a path $\Gamma : [0,1] \to \Pi_{\mathcal{E}}, \Gamma(0) = \gamma_0, \Gamma(1) = \gamma_1$ is not possible without having $\Gamma(t) \notin \mathbb{B}_{d_{FP}}(\gamma_1, d_{FP}(\gamma_0, \gamma_1))$ for some $t \in [0,1]$.

Observation 26. *Examining the path-connectivity of $\mathcal{G}_{\mathcal{E}}$ under the Fréchet distance reduces to a knot theory problem for $n \leq 3$. For $n \geq 4$, there exists a sequence of Reidemeister moves from any tame knot to another, and we can construct a path in $\mathcal{G}_{\mathcal{E}}$ by interpolating and conducting Reidemeister moves at a self-crossing event.*

Lemma 27 (Path-Connectivity of $(\mathcal{G}_{\mathcal{E}}, d_{FG})$, $n \leq 3$)**.** *In general, the topological space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ is not path-connected for a graph G, if G is embedded in $\mathbb{R}^n$ with $n \leq 3$.*

Proof. If embeddings in $\mathbb{R}^n$ are restricted to $n \leq 3$, then as a consequence of knot theory, $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ is not path-connected for any graph G.

If $n = 2$, let G denote a graph consisting of only a cycle comprising two vertices, and a single dangling edge. Let $\mathbf{G}_0 = (G, \phi_0) \in \mathcal{G}_{\mathcal{E}}$ comprise a closed curve with an interior edge, and let $\mathbf{G}_1 = (G, \phi_1) \in \mathcal{G}_{\mathcal{E}}$ comprise a closed curve with an exterior edge in $\mathbb{R}^n$. By the Jordan curve theorem, there does not exist a continuous path in $\mathbb{R}^n$ from $\mathbf{G}_0$ to $\mathbf{G}_1$ that does not create a degeneracy. Then, constructing a path from $\mathbf{G}_0$ to $\mathbf{G}_1$ must reach some juncture where an immersed graph in $\mathbb{R}^n$, denoted $\mathbf{G}_* = (G, \phi_0^*)$, is not homeomorphic to G. Therefore, $\mathbf{G}_*$ violates the definition of a graph embedding, and the space $(\mathcal{G}_{\mathcal{E}}, d_{GF})$ is not path-connected among homeomorphism classes of graphs in dimension 2.

If $n = 3$, let G consist of a single cycle, and $\mathbf{G}_0 = (G, \phi_0) \in \mathcal{G}_{\mathcal{E}} = \mathbb{S}^1$ and $\mathbf{G}_1 = (G, \phi_1) \in \mathcal{G}_{\mathcal{E}}$ comprise a trefoil knot. Then, again due to elementary knot theory, there exists no continuous path from $\mathbf{G}_0$ to $\mathbf{G}_1$ in the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$. See Figure 8. ☐

Theorem 28 (Path-Connectivity of $(\mathcal{G}_{\mathcal{E}}, d_{FG})$, $n \geq 4$)**.** *The topological space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ is path-connected for a graph G, if G is embedded in $\mathbb{R}^n$ with $n \geq 4$. Moreover, the connected components of the extended metric space $(\mathcal{G}_{\mathcal{E}}, d_{FG})$ are in one-to-one correspondence with the homeomorphism classes of graphs.*

Proof. Let G be a graph, and $\mathbf{G}_0 = (G, \phi_0), \mathbf{G}_1 = (G, \phi_1) \in \mathcal{G}_{\mathcal{E}}$. In dimension 4 or higher, it is well known that any tame knot can be unwound by a sequence of Reidemeister moves into the unknot. Then, one may interpolate along the pointwise leashes defining $d_{FG}(\mathbf{G}_0, \mathbf{G}_1)$ until a crossing event must occur. At this juncture, there must exist a Reidemeister move allowing the crossing event to occur. Hence, any sequence of knots and dangling edges comprising the image of ϕ_0 can be unwound to a sequence of unknots and straight edges. The same holds for the image of ϕ_1. Consequently there exists a continuous path from $\mathbf{G}_0$ to $\mathbf{G}_1$ in the topological space $\mathcal{G}_{\mathcal{E}}(G, d_{FG})$. Note that we require that ϕ_0, ϕ_1 are rectifiable in Section 2. Without this requirement, $\mathbf{G}_0$ and $\mathbf{G}_1$ could comprise wild knots, and constructing such a path could consist of infinitely many Reidemeister moves. □

Theorem 29 (Metric Balls in $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$). *Metric balls in the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ are path-connected if $n \geq 4$*

Proof. The proof is identical to that in Lemma 37 but considers graph embeddings, and is due to the path constructed in Lemma 38. □

5 Conclusion

In this paper, we established fundamental topological properties of generalized spaces of paths and graphs in Euclidean space under the Fréchet distance. We provided proof of the path-connectedness property, and gave tight bounds in lower dimensional settings that are not path-connected under the Fréchet distance. Our results in low dimensiona show that the Fréchet distance can be problematic for the well-definedness of spaces of paths and graphs in $\mathbb{R}^n$ after perturbation. This could arise, for example, in nearest neighbor models and more broadly in machine learning. We additionally investigated metric properties of the Fréchet distance for graphs, and made stronger guarantees about the path-connectivity of metric balls in these spaces. Due to the widespread popularity of the Fréchet distance in computational geometry and growing interest in the machine learning community, establishing the underlying properties of the topological spaces it can define sets an important theoretical backdrop. Our contribution establishes these notions regarding the Fréchet distance for the first time, and lays the theoretical foundation for the extension of the Fréchet distance to new problems in machine learning related to graph-based data. Extensions to this work abound, and include examining core topological properties of other distance measures in computational geometry.

Acknowledgements

References

[1] P. K. Agarwal, R. B. Avraham, H. Kaplan, and M. Sharir. Computing the discrete Fréchet distance in subquadratic time. *SIAM Journal on Computing*, 43(2):429–449, 2014.

[2] M. Ahmed and C. Wenk. Constructing street networks from gps trajectories. In *Algorithms – ESA 2012*, pages 60–71, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[3] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *IJCGA*, 5(1–2):75–91, 1995.

[4] H. Alt, C. Knauer, and C. Wenk. Matching polygonal curves with respect to the Fréchet distance. In A. Ferreira and H. Reichel, editors, *STACS 2001*, pages 63–74, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

[5] J. Anjaria, H. Wei, H. Li, S. Mishra, and H. Sa, et. Trajdistlearn: learning to compute distance between trajectories. In *Proceedings of the 14th ACM SIGSPATIAL International Workshop on Computational Transportation Science (IWCTS '21)*, volume 4, pages 1–9, 2021.

[6] B. Aronov, S. Har-Peled, C. Knauer, Y. Wang, and C. Wenk. Fréchet distance for curves, revisited. In Y. Azar and T. Erlebach, editors, *Algorithms – ESA 2006*, pages 52–63, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[7] K. Buchin, M. Buchin, and A. Schulz. Fréchet distance of surfaces: Some simple hard cases. In *European Symposium on Algorithms*, pages 63–74. Springer, 2010.

[8] K. Buchin, T. Ophelders, and B. Speckmann. Computing the Fréchet distance between real-valued surfaces. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2443–2455. ACM, 2017.

[9] M. Buchin, A. Krivosija, and A. Neuhaus. Computing the Fréchet distance of trees and graphs of bounded tree width. In *Proceedings of the 36th European Workshop on Computational Geometry*, 2020.

[10] E. Chambers, B. T. Fasy, Y. Wang, and C. Wenk. Map-matching using shortest paths. In *ACM Transactions on Spatial Algorithms and Systems*, pages 1–17. Association for Computing Machinery, 2020.

[11] E. W. Chambers, E. Colin de Verdière, J. Erickson, S. Lazard, F. Lazarus, and S. Thite. Homotopic Fréchet distance between curves or, walking your dog in the woods in polynomial time. *Computational Geometry*, 43(3):295–311, 2010. Special Issue on 24th Annual Symposium on Computational Geometry (SoCG'08).

[12] D. Chen, A. Driemel, L. J. Guibas, A. Nguyen, and C. Wenk. Approximate map matching with respect to the fr ́echet distance. pages 75–83, 2011.

[13] C. Colombe and K. Fox. Approximating the (continuous) fréchet distance. In K. Buchin and E. Colin de Verdiére, editors, *37th International Symposium on Computational Geometry (SoCG, 2021)*, 2021.

[14] A. Driemel and S. Har-Peled. Jaywalking your dog: Computing the Fréchet distance with shortcuts. *SIAM Journal on Computing*, 42(5):1830–1866, 2013.

[15] A. Driemel, S. Har-Peled, and C. Wenk. Approximating the Fréchet distance for realistic curves in near linear time. *Discrete and Computational Geometry*, 48(1):94–127, Feb. 2012.

[16] A. Driemel, I. van der Hoog, and E. Rotenburg. On the discrete fréchet distance in a graph. In M. Kerber and X. Goaoc, editors, *Proceedings of the Symposium on Computational Geometry, 2022*.

[17] H. Edelsbrunner and J. L. Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010.

[18] P. Fang and C. Wenk. The Fréchet distance for plane graphs. In *Proceedings of the 37th European Workshop on Computational Geometry*, 2021.

[19] A. Filtser, O. Filtser, and M. J. Katz. Approximate Nearest Neighbor for Curves - Simple, Efficient, and Deterministic. In A. Czumaj, A. Dawar, and E. Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:19, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[20] M. Fréchet. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Mathematico di Palermo*, 22:1–74, 1906. Paragraphs 78–80.

[21] P. Indyk. Approximate nearest neighbor algorithms for frechet distance via product metrics. In *SCG '02: Proceedings of the Eighteenth Annual Symposium on Computational Geometry*, pages 102–106, 2002.

[22] A. F. C. IV and C. Wenk. Geodesic Fréchet distance inside a simple polygon. *ACM Trans. Algorithms*, 7(1), dec 2010.

[23] M. Musleh, S. Abbar, R. Stanojevic, and M. Mokbel. Qarta: an ml-based system for accurate map services. *Proceedings of the VLDB Endowment*, 14(11):2273—2282, July 2021.

[24] M. T. Pilehvar and J. Camacho-Collados. *Embeddings in Natural Language Processing*. Morgan and Claypool Publishers, 2020.

[25] M. Xu. Understanding graph embedding methods and their applications. *SIAM Review*, 63(4):825–853, 2021.

In what follows, we include the smaller results and clarifications omitted for brevity throughout the paper.

A Additional Details for Section 2

A.1 Formulations of Spaces of Graphs from Dimensionality Reduction

Techniques throughout the machine learning literature [25, 24] discuss a large number of dimensionality reduction methods, which can be understood as a map $f : G \to \mathbb{R}^n$ taking a graph G to a lower dimensional Euclidean space. In doing so, f generally produces an embedding of just the vertices, which preserves graph-based distances between vertices with Euclidean distances. Such mappings typically disregard the images of edges in $\mathbb{R}^n$. However, assuming that f requires that the *vertices* are indeed embedded in $\mathbb{R}^n$, i.e. all vertices are mapped to unique coordinates, there is a canonical graph immersion mapping all edges $e = \{v_1, v_2\}$ in G to the straight line between v_1 and v_2. Such a map is vacuously an immersion, because no requirement for injectivity of f is set, but local injectivity must be preserved in order for f to be injective among the vertices. Hence, in these settings it is reasonable to expect that graph "embeddings" in the traditional context of dimensionality reduction are operating on graph immersions in our definition, within the set $\mathcal{G}_{\mathcal{I}}(G)$.

B Additional Details for Section 4

B.1 Additional Details for Section 4.1

Given two continuous mappings of the same graph into $\mathbb{R}^n$, we can interpolate between them. First, we need to define linear combinations of graphs (and paths).

Definition 30 (Linear Combination of Graphs)**.** *Let G be a graph. Let $\phi_0 \colon G \to \mathbb{R}^n$ and $\phi_1 \colon G \to \mathbb{R}^n$ be continuous, and denote $\mathbf{G}_0 = (G, \phi_0), \mathbf{G}_1 = (G, \phi_1) \in \mathcal{G}_{\mathcal{C}}(G)$. If $h \colon G \to G$ is a homeomorphism and $c_0, c_1 \in \mathbb{R}$, then the linear combination $c_0\mathbf{G}_0 + c_1\mathbf{G}_1$ with respect to h is defined as follows: we define $\phi \colon G \to \mathbb{R}^n$ by $\phi(x) := c_0\phi_0(x) + c_1\phi_1(h(x))$. Abusing notation, we write $c_0\mathbf{G}_0 + c_1\mathbf{G}_1 = (G, \phi)$ as this linear combination with respect to h, the homeomorphism defining $d_{FG}(\mathbf{G}_0, \mathbf{G}_1)$.*

Above, we observe that ϕ is continuous (since ϕ_0 and ϕ_1 are continuous). In addition, we note that linear combinations of graphs are defined *on the specific representations* of the continuously mapped graphs, not *on the elements of $\mathcal{G}_{\mathcal{C}}$*. For a simple example taking $G = [0, 1]$, see Figure 5.

Definition 31 (Linear Interpolation)**.** *Let G be a graph and $\mathbf{G}_0, \mathbf{G}_1 \in \mathcal{G}_{\mathcal{C}}(G)$. We define the* linear interpolation *from $\mathbf{G}_0$ to $\mathbf{G}_1$ to be the function $\Gamma \colon [0,1] \to$*

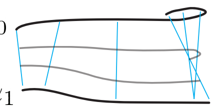


Figure 5: The interpolation from one path to another, by linear combinations defined in Definition 30.

$\mathcal{G}_\mathcal{C}(G)$, where:

$$\Gamma(t) := (1-t)\mathbf{G}_0 + t\mathbf{G}_1. \tag{2}$$

That is, Γ is a continuous family of linear combinations of the graphs $\mathbf{G}_0$ and $\mathbf{G}_1$. In fact, (2) is the linear combination defined in Definition 30, where we take $c_0 = 1-t$ and $c_1 = t$ from the definition. We note that if $G = [0,1]$, the linear interpolation among graphs is nothing other than linear interpolation between paths.

The fact that Γ is well-defined and continuous is proven in the next lemma.

Lemma 32 (Properties of Interpolation)**.** *Linear interpolation $\Gamma : [0,1] \to \mathcal{G}_\mathcal{C}(G)$ between graphs in $\Pi_\mathcal{C}$ and $\mathcal{G}_\mathcal{C}(G)$ is well-defined and is a continuous function.*

Proof. We demonstrate that linear interpolation as given in Definition 31 is well defined and continuous. From Definition 30, at any $t \in [0,1]$, $\Gamma(t)$ is well defined in the space $\mathcal{G}_\mathcal{C}$ of continuous mappings since any such linear combination of graphs represents a continuous map from the underlying graph to $\mathcal{G}_\mathcal{C}$.

To see that Γ is continuous, we will show that the preimage of any open set S in $\mathcal{G}_\mathcal{C}(G)$ is open after applying Γ. Without loss of generality, we can take $S \subseteq \mathcal{G}_\mathcal{C}(G) \cap Image(\Gamma))$, since any $\mathbf{G} \in \mathcal{G}_\mathcal{C}(G) \notin Image(\Gamma)$ has $\Gamma^{-1} = \emptyset$. Examine an open Fréchet ball in $\mathcal{G}_\mathcal{C}(G) \cap Image(\Gamma)$ of radius $r = \delta/d_{FG}(\mathbf{G}_0, \mathbf{G}_1)$ centered at a graph mapping $\mathbf{G}_t$ in the image of Γ for some $\delta \in \mathbb{R}$ such that $d_{FG}(\mathbf{G}_0, \mathbf{G}_1) > \delta > 0$. This is denoted $\mathbb{B}_{\mathcal{G}_\mathcal{C}(G)}(\mathbf{G}_t, r) \subset Im(\Gamma)$, where $\Gamma(t) = \mathbf{G}_t$. Observe that

$$\Gamma^{-1}(\mathbb{B}_{\mathcal{G}_\mathcal{C}(G)}(\mathbf{G}_t, r)) = (t-\delta, t+\delta) \subset [0,1]$$

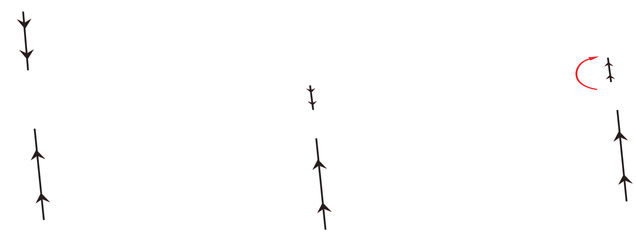
which is an open interval.

Moreover, construct an arbitrary open set $S \subseteq \mathcal{G}_\mathcal{C}(G)$ by taking the (possibly infinite) union of open Fréchet balls defined in an identical manner to the one above:

$$S := \bigcup_{i=1}^{\infty} \{B_i : B_i = \mathbb{B}_{\mathcal{G}_\mathcal{C}(G)}(\mathbf{G}_{t_i}, r_i)\}$$

for any desired $t_i \in [0,1], r_i = \delta_i/d_{FG}(\mathbf{G}_0, \mathbf{G}_1), \delta_i > 0$.

By design, $\Gamma^{-1}(S) = \bigcup_{i=1}^{\infty}\{(t_i - \delta_i, t_i + \delta_i)\} \subseteq [0,1]$, a (possibly infinite) union of open intervals within $[0,1]$. Hence, since Γ^{-1} acting on any open set is



(a) Paths with reversed orientation

(b) Interpolate

(c) Rotate when sufficiently close

Figure 6: For colinear paths with opposing orientation, rotating by π will avoid degenerating to the constant map, keeping Γ in $\Pi_{\mathcal{I}}$. Moreover, rotation with sufficiently small Fréchet distance maintains the path-connectivity of balls.

open by design, Γ is continuous. Finally, we can conclude that $(\mathcal{G}_{\mathcal{C}}(G), d_{FG})$ is path-connected.

Setting $G = [0,1]$, we obtain that linear interpolation between paths in $\Pi_{\mathcal{C}}$ is a well-defined continuous function in an identical manner. $\square$

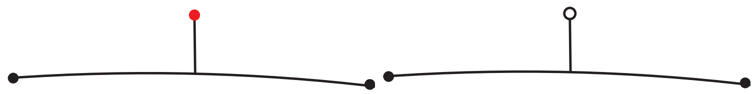
B.2 Additional Details for Section 4.2

Lemma 33 (Dodging Singletons)**.** *Suppose at a time $t \in [0,1]$, a linear interpolation $\Gamma(t)$ in $\Pi_{\mathcal{I}}$ would result in a constant map γ_t. This total degeneracy can be averted by rotating $\Gamma(t)$.*

Proof. Linear interpolation of γ_0 to γ_1 can only produce a singleton if the two paths are colinear with reversed orientation. Hence, if $\Gamma(t)$ degenerates to a constant map, there exists sufficiently small $\epsilon > 0$ such that $\Gamma(t-\epsilon)$ can be continuously rotated by π without forcing $d_{FP}(\gamma_0, \gamma_1) < d_{FP}(\Gamma(t), \gamma_1)$. Thereby reversing the orientation of $\Gamma(t+\epsilon)$, and avoiding the constant map at $\Gamma(t)$. See Figure 6 for an example. $\square$

Lemma 34 (The Q-Tip Maneuver)**.** *Suppose at a time $t \in [0,1]$, a linear interpolation $\Gamma(t)$ in $\Pi_{\mathcal{I}}$ would result in backtracking within γ_t. This violation of injectivity can be corrected by inflating a ball about the critical backtracking point.*

Proof. In the scenario of a backtracking event, local injectivity is only violated at the exact critical point $\gamma_t(x)$ for $x \in [0,1]$ where backtracking occurs. Denote $\gamma_{t-\epsilon} = \Gamma(t-\epsilon)$. For sufficiently small $\epsilon, \delta > 0$, continuously inflate a ball of radius δ about $\gamma(x)$ such that $d_{FP}(\gamma_{t-\epsilon}, \gamma_1)$ remains fixed, completing at time t with a ball replacing the degeneracy, and maintaining that $\gamma_t \in \Pi_{\mathcal{I}}$. For an example of this maneuver, see Figure 7b. $\square$



(a) Example path with backtracking (b) Inflate the critical backtracking point

Figure 7: Reconcile forced backtracking along a path by inflating a ball about the critical backtracking point, thereby maintaining local injectivity.

B.3 Additional Details for Section 4.3

Lemma 35. *Metric balls in the space* $(\Pi_{\mathcal{E}}, d_{FP})$ *of embedded paths in* $\mathbb{R}^n$ *are not path-connected if* $n = 2$

Proof. Metric balls are not in general path-connected in two dimensions. For example, suppose γ_0 comprises a segment of large width and small height in $\mathbb{R}^2$, and γ_1 comprises its mirror image. Suppose the height of γ_0 is $\delta/2$, the width of γ_0 is 2δ, and the distance $d_{FP}(\gamma_0, \gamma_1) = \delta$, defined by the endpoints of γ_0 and γ_1. Then any path from γ_0 to γ_1 must flip an endpoint across the segment, necessarily increasing the Fréchet distance. See Figure 4. □

Lemma 36. *Metric balls in the space* $(\Pi_{\mathcal{E}}, d_{FP})$ *of embedded paths in* $\mathbb{R}^n$ *are not path-connected if* $n = 3$

Proof. Metric balls are not in general path-connected in three dimensions. For a simple counterexample, suppose γ_0 comprises a loop in $\mathbb{R}^3$, where a segment crossed on top of itself, avoiding self-intersection by some small distance δ, with long tails at either end of the crossing of length 2δ. Suppose also that γ_1 comprises the mirror image of γ_0. Then $d_{FP} = \delta$, but it is not possible to construct a path from γ_0 to γ_1 without increasing the Fréchet distance between the two, since γ_0 must conduct a self-crossing, which will increase the Fréchet distance by at least 2δ. Again, see Figure 4 □

Lemma 37 (Metric Balls in $(\Pi_{\mathcal{E}}, d_{FP})$, $n \geq 4$)**.** *Metric balls in the space of paths embedded in* $\mathbb{R}^n$ *topologized by the Fréchet distance* $(\Pi_{\mathcal{E}}, d_{FP})$ *are path-connected for* $n \geq 4$.

Proof. Let $\gamma_0, \gamma_1 \in \Pi_{\mathcal{E}}$ in the ambient space $\mathbb{R}^n$, for $n \geq 4$. Since all topological knots are represented equivalently in only 3 dimensions, we can consider without loss of generality the projections of γ_0 and γ_1 in $\mathbb{R}^3$. Suppose $d_{FP}(\gamma_0, \gamma_1) = \delta$. Then, construct a continuous $\Gamma : [0, 1] \to \Pi_{\mathcal{E}}$ be the linear interpolation from $\Gamma(0) = \gamma_0$ to $\Gamma(1) = \gamma_1$. By the rectifiability of the embeddings γ_0 and γ_1, the interpolation must reduce $d_{FP}(\gamma_0, \gamma_1)$ by some $\epsilon > 0$ before a self-crossing is required in the image γ_t of Γ at time $t \in [0, 1]$.

At this time t, conduct a self-crossing by perturbing $\Gamma(t)$ in the fourth dimension by no more than $\epsilon/2$. This will increase $d_{FP}(\gamma_t, \gamma_1)$ by no more than $\epsilon/2$. Hence, $d_{FP}(\gamma_t, \gamma_1)$ is either strictly decreasing as $t \to 1$, or necessarily satisfies

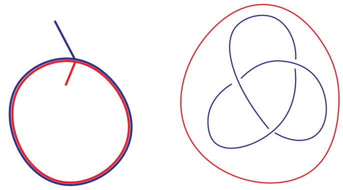


Figure 8: Embedded graphs in $\mathbb{R}^2$ and $\mathbb{R}^3$ which are not path-connected in the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$.

$d_{FP}(\gamma_t, \gamma_1) \leq \delta - \epsilon/2$ for $\epsilon > 0$. This is to say, for all $t \in [0, 1]$, $d_{FP}(\gamma_t, \gamma_1) \leq \delta$, and metric balls in the space are path-connected. □

Lemma 38. *Let G be a graph. Metric balls in the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ of graph embeddings in $\mathbb{R}^n$ are path-connected if $n \geq 4$, which is a tight bound.*

Proof. Let $\mathbf{G}_0, \mathbf{G}_1 \in \mathcal{G}_{\mathcal{E}}$ and suppose $d_{FG}(\mathbf{G}_0, \mathbf{G}_1) = \delta$. If $d < 4$, the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ is not path-connected, which implies that metric balls are not path-connected.

However, if $n \geq 4$, metric balls in the space $(\mathcal{G}_{\mathcal{E}}(G), d_{FG})$ are path-connnected. The proof is identical to that in Lemma 37, but extended to graphs. Since every topological knot is represented in only 3 dimensions, we again consider without loss of generality the projections of $\mathbf{G}_0, \mathbf{G}_1$ in $\mathbb{R}^3$, and conduct a direct linear interpolation. When a self-cross event occurs at $\Gamma(t)$ for some $t \in (0, 1)$, the idea is identical to Lemma 37. If $d_{FG} = \delta - \epsilon$, we perturb $\Gamma(t)$ by $\epsilon/2$ in the fourth dimension and then directly conduct the crossing. We then repeat the process of interpolating and perturbing in the fourth dimension for all subsequent self-crossings. □