EXPLICIT GEODESICS IN GROMOV-HAUSDORFF SPACE

SAMIR CHOWDHURY AND FACUNDO MÉMOLI

ABSTRACT. We provide an alternative, constructive proof that the collection \mathcal{M} of isometry classes of compact metric spaces endowed with the Gromov-Hausdorff distance is a geodesic space. The core of our proof is a construction of explicit geodesics on \mathcal{M} . We also provide several interesting examples of geodesics on \mathcal{M} , including a geodesic between \mathbb{S}^0 and \mathbb{S}^n for any $n \geq 1$.

1. Geodesics on Gromov-Hausdorff space

The collection of compact metric spaces, denoted \mathcal{M} throughout this paper, is a valid pseudometric space when endowed with the Gromov-Hausdorff distance [4, 3]. We will denote this space by $(\mathcal{M}, d_{\mathrm{GH}})$. Furthermore, the space $(\mathcal{M}/\sim, d_{\mathrm{GH}})$, where we define

$$(X, d_X) \sim (Y, d_Y) \iff (X, d_X)$$
 is isometric to (Y, d_Y) , and $d_{GH}([X], [Y]) := d_{GH}(X, Y)$,

is a metric space [3]. It is known that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is separable, complete [8], and geodesic [6]. Specifically, the authors of [6] use a compactness result to argue that for any pair of points in \mathcal{M}/\sim , there exists a midpoint in \mathcal{M}/\sim , which implies that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is a geodesic space [3, Theorem 2.4.16]. However, this proof is not constructive. The goal of our paper is to provide a constructive proof through the explicit description of a certain class of geodesics on $(\mathcal{M}/\sim, d_{\mathrm{GH}})$, which we call straight-line geodesics. The key ingredient in our construction is a proof showing that there exists an optimal correspondence between any two compact metric spaces. While obvious when considering finite metric spaces, establishing this result for general compact metric spaces requires some work. Our result is inspired by a similar result proved by Sturm about geodesics on the space of metric measure spaces [9]. We use our result to construct: (1) an explicit geodesic between \mathbb{S}^0 and \mathbb{S}^n , for any $n \in \mathbb{N}$, and (2) explicit, infinite families of deviant (i.e. non-straight-line) and branching geodesics between the one-point discrete space and the n-point discrete space, for any $n \geq 2$.

Before proceeding, we recall some concepts, in particular that of a geodesic space. A curve in a metric space (X, d_X) is a continuous map $\gamma : [0, 1] \to X$, and its length is given by:

$$L(\gamma) := \sup \left\{ \sum_{i=1}^{n-1} d_X(\gamma(t_i), \gamma(t_{i+1}) : 0 = t_1 \le t_2 \le \dots \le t_n = 1, n \in \mathbb{N} \right\}.$$

²⁰¹⁰ Mathematics Subject Classification. Primary 53C23, Secondary 51F99. This work was supported by NSF grants CCF-1526513 and IIS-1422400.

Such a curve is called a *geodesic* [2, Section I.1] if for any $s, t \in [0, 1]$,

$$d_X(\gamma(s), \gamma(t)) = |t - s| \cdot d_X(\gamma(0), \gamma(1)).$$

As a consequence of this definition, for any geodesic γ such that $\gamma(0) = x$ and $\gamma(1) = x'$, one has $L(\gamma) = d_X(x, x')$. The metric space (X, d_X) is called a *geodesic space* if for any $x, x' \in X$, there exists a geodesic γ connecting x and x'.

Next, given a metric space (X, d_X) and two nonempty subsets $A, B \subseteq X$, the *Hausdorff distance* between A and B is defined as:

$$d_{\mathrm{H}}^{X}(A,B) := \max \left(\sup_{a \in A} \inf_{b \in B} d_{X}(a,b), \sup_{b \in B} \inf_{a \in A} d_{X}(a,b) \right).$$

Given $(X, d_X), (Y, d_Y) \in \mathcal{M}$, the Gromov-Hausdorff distance between them is defined as:

(1.1)
$$d_{\mathrm{GH}}((X, d_X), (Y, d_Y)) := \inf_{\substack{(Z, d_Z) \in \mathcal{M} \\ \varphi: X \to Z, \ \psi: Y \to Z}} d_{\mathrm{H}}^Z(\varphi(X), \psi(Y)),$$

where φ and ψ are both isometric embeddings [3]. Notice that d_{GH} is well-defined on $[X], [Y] \in \mathcal{M}/\sim$. Indeed, if $X' \in [X], Y' \in [Y]$, then:

$$d_{GH}([X'], [Y']) = d_{GH}(X', Y') = d_{GH}(X, Y) = d_{GH}([X], [Y]),$$

where the second-to-last equality follows from the triangle inequality and the observation that $d_{GH}(X, X') = d_{GH}(Y, Y') = 0$.

It is known that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is complete [8]. Details about the topology generated by the Gromov-Hausdorff distance can be found in [3]. One important fact is that it allows the existence of many compact sets in $(\mathcal{M}/\sim, d_{\mathrm{GH}})$, in the sense below. Recall that for a compact metric space X, for $\varepsilon > 0$, the ε -covering number $\operatorname{cov}_X(\varepsilon)$ is defined to be the minimum number of ε -balls required to cover X.

Theorem 1.1 (Gromov's precompactness theorem, [8]). Given a bounded function $N:(0,\infty)\to\mathbb{N}$ and D>0, let $\mathcal{C}(N,D)\subseteq (\mathcal{M}/\sim,d_{\mathrm{GH}})$ be the collection of all [X] such that $\mathrm{diam}(X)< D$ and $\mathrm{cov}_X(\varepsilon)\leq N(\varepsilon)$ for each $\varepsilon>0$. Then $\mathcal{C}(N,D)$ is precompact.

In our constructions, we will use an equivalent formulation of the Gromov-Hausdorff distance following [3, Chapter 7]. Given $(X,d_X),(Y,d_Y)\in\mathcal{M}$, we say that a relation $R\subseteq X\times Y$ is a correspondence if for any $x\in X$, there exists $y\in Y$ such that $(x,y)\in R$, and for any $y\in Y$, there exists $x\in X$ such that $(x,y)\in R$. The set of all such correspondences will be denoted $\mathscr{R}(X,Y)$. In the case Y=X, a particularly useful correspondence is the diagonal correspondence $A:=\{(x,x):x\in X\}$. The distortion of any non-empty relation $R\subset X\times Y$ is defined to be:

$$dis(R) := \sup_{(x,y),(x',y')\in R} |d_X(x,x') - d_Y(y,y')|.$$

The Gromov-Hausdorff distance $d_{GH}: \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+$ can be formulated as:

$$d_{\mathrm{GH}}((X,d_X),(Y,d_Y)) := \frac{1}{2} \inf_{R \in \mathscr{R}(X,Y)} \mathrm{dis}(R).$$

In particular, a correspondence is *optimal* if the infimum is achieved. We will denote by $\mathscr{R}^{\mathrm{opt}}(X,Y)$ the set of all closed optimal correspondences. We have:

Proposition 1.1. $\mathscr{R}^{\mathrm{opt}}(X,Y) \neq \emptyset$ for any (X,d_X) and $(Y,d_Y) \in \mathcal{M}$.

Our main result is the explicit construction of straight-line geodesics:

Theorem 1.2 (Existence of straight-line geodesics). $(\mathcal{M}/\sim, d_{GH})$ is a geodesic space. More specifically, let $[X], [Y] \in (\mathcal{M}/\sim, d_{GH})$. Then, for any $R \in \mathscr{R}^{\text{opt}}(X, Y)$, we can construct a geodesic $\gamma_R : [0, 1] \to \mathcal{M}/\sim$ between [X] and [Y] as follows:

$$\gamma_R(0) := [(X, d_X)], \ \gamma_R(1) := [(Y, d_Y)], \ and \ \gamma_R(t) := [(R, d_{\gamma_R(t)})] \ for \ t \in (0, 1),$$

where for each $(x, y), (x', y') \in R$ and $t \in (0, 1)$,

$$d_{\gamma_R(t)}((x,y),(x',y')) := (1-t) \cdot d_X(x,x') + t \cdot d_Y(y,y').$$

Not all geodesics between compact metric spaces are of the form given by Theorem 1.2. Furthermore, *branching* of geodesics may happen in Gromov-Hausdorff space. We explore the deviance and branching phenomena in Section 1.1. In Section 1.2 we construct explicit geodesics between \mathbb{S}^0 and \mathbb{S}^n . The proofs of Proposition 1.1 and Theorem 1.2 are given in Section 2.

1.1. **Deviant and branching geodesics.** The following lemma will be useful in the sequel:

Lemma 1.3. Let (Z, d_Z) be a metric space. Let $S, T \in \mathbb{R}$, with S < T, and $\gamma : [S, T] \to Z$ be a curve such that

$$d_Z(\gamma(s), \gamma(t)) \le \frac{|s-t|}{|S-T|} \cdot d_Z(\gamma(S), \gamma(T)), \text{ for all } s, t \in [S, T].$$

Then, in fact,

$$d_Z(\gamma(s), \gamma(t)) = \frac{|s-t|}{|S-T|} \cdot d_Z(\gamma(S), \gamma(T)), \text{ for all } s, t \in [S, T].$$

Proof of Lemma 1.3. Suppose the inequality is strict. Suppose also that $s \leq t$. Then by the triangle inequality, we obtain:

$$d_Z(\gamma(S), \gamma(T)) \le d_Z(\gamma(S), \gamma(s)) + d_Z(\gamma(s), \gamma(t)) + d_Z(\gamma(t), \gamma(T))$$

$$< \frac{(s-S) + (t-s) + (T-t)}{T-S} \cdot d_Z(\gamma(S), \gamma(T)).$$

This is a contradiction. Similarly we get a contradiction for the case t < s. This proves the lemma.

1.1.1. Deviant geodesics. For any $n \in \mathbb{N}$, let Δ_n denote the n-point discrete space, often called the n-point unit simplex. Fix $n \in \mathbb{N}$, $n \geq 2$. We will construct an infinite family of deviant geodesics between Δ_1 and Δ_n , named as such because they deviate from the straight-line geodesics given by Theorem 1.2. As a preliminary step, we describe the straight-line geodesic between Δ_1 and Δ_n of the form given by Theorem 1.2. Let $\{p\}$ and $\{x_1,\ldots,x_n\}$ denote the underlying sets of Δ_1 and Δ_n . There is a unique correspondence $R := \{(p,x_1),\ldots,(p,x_n)\}$ between these two sets. According to the setup in Theorem 1.2, the straight-line geodesic between Δ_1 and Δ_n is then given by the metric spaces $(R,d_{\gamma_R(t)})$, for $t \in (0,1)$. Here $d_{\gamma_R(t)}((p,x_i),(p,x_j)) = t \cdot d_{\Delta_n}(x_i,x_j) = t$ for each $t \in (0,1)$ and each $1 \leq i,j \leq n$. This corresponds to the all-t matrix with 0s on the diagonal. Finally, we note that the unique correspondence R necessarily has distortion 1. Thus $d_{\mathrm{GH}}(\Delta_1,\Delta_n) = \frac{1}{2}$.

Now we give the parameters for the construction of a certain family of deviant geodesics between Δ_1 and Δ_n . For any $\alpha \in (0,1]$ and $t \in [0,1]$, define

$$f(\alpha, t) := \begin{cases} t\alpha &: 0 \le t \le \frac{1}{2} \\ \alpha - t\alpha &: \frac{1}{2} < t \le 1 \end{cases}$$

Next let m be a positive integer such that $1 \leq m \leq n$, and fix a set $X_{n+m} := \{x_1, x_2, \ldots, x_{n+m}\}$. Fix $\alpha_1, \ldots, \alpha_m \in (0, 1]$. For each $0 \leq t \leq 1$, define the matrix $\delta_t := ((d_{ij}^t))_{i,j=1}^{n+m}$ by:

For
$$1 \le i, j \le n + m$$
, $d_{ij}^t := \begin{cases} 0 & : i = j \\ f(\alpha_i, t) & : j - i = n \\ f(\alpha_j, t) & : i - j = n \\ t & : \text{ otherwise.} \end{cases}$

This is a block matrix $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ where A is the $n \times n$ all-t matrix with 0s on the diagonal, C is an $m \times m$ all-t matrix with 0s on the diagonal, and B is the $n \times m$ all-t matrix with $f(\alpha_1, t), f(\alpha_2, t), \ldots, f(\alpha_m, t)$ on the diagonal.

We first claim that δ_t is the distance matrix of a pseudometric space. Symmetry is clear. We now check the triangle inequality. In the cases $1 \leq i, j, k \leq n$ and $n+1 \leq i, j, k \leq n+m$, the points x_i, x_j, x_k form the vertices of an equilateral triangle with side length t. Suppose $1 \leq i, j \leq n$ and $n+1 \leq k \leq n+m$. Then the triple x_i, x_j, x_k forms an isosceles triangle with equal longest sides of length t, and a possibly shorter side of length $f(\alpha_i, t)$ (if |k-i|=n), $f(\alpha_j, t)$ (if |k-j|=n), or just a third equal side with length t in the remaining cases. The case $1 \leq i \leq n$, $n+1 \leq j, k \leq n+m$ is similar. This verifies the triangle inequality. Also note that δ_t is the distance matrix of a bona fide metric space for $t \in (0,1)$. For t=1, we identify the points x_i and x_{i-n} , for $n+1 \leq i \leq n+m$, to obtain Δ_n , and for t=0, we identify all points together to obtain Δ_1 . This allows us to define geodesics between Δ_1 and Δ_n as follows. Let $\vec{\alpha}$ denote the vector $(\alpha_1, \ldots, \alpha_m)$. We define a map $\gamma_{\vec{\alpha}} : [0,1] \to \mathcal{M}$ by writing:

$$\gamma_{\vec{\alpha}}(t) := (X_{n+m}, \delta_t) \qquad t \in [0, 1],$$

where we can take quotients at the endpoints as described above.

We now verify that these curves are indeed geodesics. There are three cases: $s,t \in [0,\frac{1}{2}],\ s,t \in (\frac{1}{2},1],\$ and $s \in [0,\frac{1}{2}],\ t \in (\frac{1}{2},1].$ By using the diagonal correspondence \blacktriangle , we check case-by-case that $\mathrm{dis}(\blacktriangle) \leq |t-s|$. Thus for any $s,t \in [0,1],$ we have $d_{\mathrm{GH}}(\gamma_{\vec{\alpha}}(s),\gamma_{\vec{\alpha}}(t)) \leq \frac{1}{2}|t-s| = |t-s| \cdot d_{\mathrm{GH}}(\Delta_1,\Delta_n)$. It follows by Lemma 1.3 that $\gamma_{\vec{\alpha}}$ is a geodesic between Δ_1 and Δ_n . Furthermore, since $\vec{\alpha} \in (0,1]^m$ was arbitrary, this holds for any such $\vec{\alpha}$. Thus we have an infinite family of geodesics $\gamma_{\vec{\alpha}}: [0,1] \to \mathcal{M}$ from Δ_1 to Δ_n .

A priori, some of these geodesics may intersect at points other than the endpoints. By this we mean that there may exist $t \in (0,1)$ and $\vec{\alpha} \neq \vec{\beta} \in (0,1]^m$ such that $[\gamma_{\vec{\alpha}}(t)] = [\gamma_{\vec{\beta}}(t)]$ in \mathcal{M}/\sim . This is related to the branching phenomena that we describe in the next section. For now, we give an infinite subfamily of geodesics that do not intersect each other anywhere except at the endpoints. Recall that the separation of a finite metric space (X, d_X) is the smallest positive distance in X, which we denote by $\operatorname{sep}(X)$. If $\operatorname{sep}(X) < \operatorname{sep}(Y)$ for two finite metric spaces X and Y, then $d_{\operatorname{GH}}(X,Y) > 0$.

Let \prec denote the following relation on $(0,1]^m$: for $\vec{\alpha}, \vec{\beta} \in (0,1]^m$, set $\vec{\alpha} \prec \vec{\beta}$ if $\alpha_i < \beta_i$ for each $1 \leq i \leq m$. Next let $\vec{\alpha}, \vec{\beta} \in (0,1]^m$ be such that $\vec{\alpha} \prec \vec{\beta}$. Then $\gamma_{\vec{\beta}}$ is a geodesic from Δ_1 to Δ_n which is distinct (i.e. non-isometric) from $\gamma_{\vec{\alpha}}$ everywhere except at its endpoints. This is because the condition $\vec{\alpha} \prec \vec{\beta}$ guarantees that for each $t \in (0,1)$, $\text{sep}(\gamma_{\vec{\alpha}}(t)) < \text{sep}(\gamma_{\vec{\beta}}(t))$. Hence $d_{\text{GH}}(\gamma_{\vec{\alpha}}(t), \gamma_{\vec{\beta}}(t)) > 0$ for all $t \in (0,1)$.

Finally, let $\vec{\alpha} \in (0,1)^m$, and let $\vec{1}$ denote the all-ones vector of length m. For $\eta \in [0,1]$, define $\vec{\beta}(\eta) := (1-\eta)\vec{a} + \eta \vec{1}$. Then by the observations about the relation \prec , $\{\gamma_{\vec{\beta}(\eta)} : \eta \in [0,1]\}$ is an infinite family of geodesics from Δ_1 to Δ_n that do not intersect pairwise anywhere except at the endpoints.

Note that one could choose the diameter of Δ_n to be arbitrarily small and still obtain deviant geodesics via the construction above.

1.1.2. Branching. The structure of d_{GH} permits branching geodesics, as illustrated on the right. We use the notation $(a)^+$ for any $a \in \mathbb{R}$ to denote $\max(0, a)$. As above, fix $n \in \mathbb{N}$, $n \geq 2$, and consider the straight-line geodesic between Δ_1 and Δ_n described at the beginning of Section 1.1.1. Throughout this section, we denote this geodesic by $\gamma : [0, 1] \to \mathcal{M}$. We will construct an infinite family of geodesics which branch off from γ . For convenience, we will overload notation and write, for each



 $t \in [0,1]$, the distance matrix of $\gamma(t)$ as $\gamma(t)$. Recall from above that $\gamma(t)$ is a symmetric $n \times n$ matrix with the following form:

$$\begin{pmatrix} 0 & t & t & \dots & t \\ & 0 & t & \dots & t \\ & & \ddots & \dots & \vdots \\ & & & & 0 \end{pmatrix}$$

Let $(a_i)_{i\in\mathbb{N}}$ be any sequence such that $0 < a_1 < a_2 < \ldots < 1$. For each $t \in [0,1]$, define the $(n+1) \times (n+1)$ matrix $\gamma^{(a_1)}(t)$ to be the symmetric matrix with the following upper triangular form:

For $t > a_1$, we have $d_{\text{GH}}\left(\gamma(t), \gamma^{(a_1)}(t)\right) > 0$, because any correspondence between $\gamma(t), \gamma^{(a_1)}(t)$ has distortion at least $t - a_1$. Thus $\gamma^{(a_1)}$ branches off from γ at a_1 .

The construction of $\gamma^{(a_1)}(t)$ above is a special case of a *one-point metric extension*. Such a construction involves appending an extra row and column to the distance matrix of the starting space; explicit conditions for the entries of the new row and column are stated in [7, Lemma 5.1.22]. In particular, $\gamma^{(a_1)}(t)$ above satisfies these conditions.

Procedurally, the $\gamma^{(a_1)}(t)$ construction can be generalized as follows. Let (\bullet) denote any finite subsequence of $(a_i)_{i\in\mathbb{N}}$. We also allow (\bullet) to be the empty subsequence. Let a_j denote the terminal element in this subsequence. Then for any a_k , k > j, we can construct $\gamma^{(\bullet, a_k)}$ as follows:

- (1) Take the rightmost column of $\gamma^{(\bullet)}(t)$, replace the only 0 by $(t-a_k)^+$, append a 0 at the bottom.
- (2) Append this column on the right to a copy of $\gamma^{(\bullet)}(t)$.
- (3) Append the transpose of another copy of this column to the bottom of the newly constructed matrix to make it symmetric.

The objects produced by this construction satisfy the one-point metric extension conditions [7, Lemma 5.1.22], and hence are distance matrices of pseudometric spaces. By taking the appropriate quotients, we obtain valid distance matrices. Symmetry is satisfied by definition, and the triangle inequality is satisfied because any triple of points forms an isosceles triangle with longest sides equal. We write $\Gamma^{(\bullet)}(t)$ to denote the matrix obtained from $\gamma^{(\bullet)}(t)$ after taking quotients. As an example, we obtain the following matrices after taking quotients for $\gamma^{(a_1)}(t)$ above, for $0 \le t \le a_1$ (below left) and for $a_1 < t \le 1$ (below right):

$$\begin{pmatrix} 0 & t & \dots & t \\ & 0 & \dots & t \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & t & \dots & t & t \\ & 0 & \dots & t & t \\ & & \ddots & \vdots & \vdots \\ & & & 0 & (t - a_1) \\ & & & & 0 \end{pmatrix}$$

Now let $(a_{i_j})_{j=1}^k$ be any finite subsequence of $(a_i)_{i\in\mathbb{N}}$. For notational convenience, we write $(b_i)_i$ instead of $(a_{i_j})_{j=1}^k$. $\Gamma^{(b_i)_i}$ is a curve in \mathcal{M} ; we need to check that it is moreover a geodesic.

Let $s \leq t \in [0,1]$. Then $\Gamma^{(b_i)_i}(s)$ and $\Gamma^{(b_i)_i}(t)$ are square matrices with n+p and n+q columns, respectively, for nonnegative integers p and q. It is possible that the matrix grows in size between s and t, so we have $q \geq p$. Denote the underlying point set by $\{x_1, x_2, \ldots, x_{n+p}, \ldots, x_{n+q}\}$. Then define:

$$A:=\{(x_i,x_i): 1 \leq i \leq n+p\}, \ B:=\{(x_{n+p},x_j): n+p < j \leq n+q\}, \ R:=A \cup B.$$

Here B is possibly empty. Note that R is a correspondence between $\Gamma^{(b_i)_i}(s)$ and $\Gamma^{(b_i)_i}(t)$, and by direct calculation we have $\operatorname{dis}(R) \leq |t-s|$. Hence we have $d_{\operatorname{GH}}\left(\Gamma^{(b_i)_i}(s),\Gamma^{(b_i)_i}(t)\right) \leq \frac{1}{2} \cdot |t-s| = |t-s| \cdot d_{\operatorname{GH}}(\Delta_1,\Delta_n)$. An application of Lemma 1.3 now shows that $\Gamma^{(b_i)_i}$ is a geodesic.

The finite subsequence $(b_i)_i$ of $(a_i)_{i\in\mathbb{N}}$ was arbitrary. Thus we have an infinite family of geodesics which branch off from γ . Since the increasing sequence $(a_i)_{i\in\mathbb{N}} \in (0,1)^{\mathbb{N}}$ was arbitrary, the branching could occur at arbitrarily many points along γ .

Remark 1.4. The existence of branching geodesics shows that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is not an Alexandrov space with curvature bounded below [3, Chapter 10]. Moreover, the existence of deviant (i.e. non-unique) geodesics shows that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ cannot have curvature bounded from above, i.e. $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is not a CAT(k) space for any k>0 [2, Proposition 2.11].

1.2. An explicit geodesic from \mathbb{S}^0 to \mathbb{S}^n . Let $n \in \mathbb{N}$. Consider the spheres \mathbb{S}^0 and \mathbb{S}^n equipped with the canonical geodesic metric, such that each sphere has diameter π . We will now construct an explicit geodesic between \mathbb{S}^0 and \mathbb{S}^n .

Proposition 1.2. $d_{GH}(\mathbb{S}^0, \mathbb{S}^n) = \frac{\pi}{2}$.

Proof of Proposition 1.2. Let U and L denote the closed upper hemisphere and open lower hemisphere of \mathbb{S}^n , respectively. Then we have $U \sqcup L = \mathbb{S}^n$. Moreover, let $s, s' \in U$ be two points realizing the diameter of \mathbb{S}^n via an arc in U (note that such an arc exists for each \mathbb{S}^n when $n \geq 1$). Also let $\{p, q\}$ denote the two points of \mathbb{S}^0 . Now we construct a correspondence between \mathbb{S}^0 and \mathbb{S}^n :

$$R := \{(p, u) : u \in U\} \cup \{(q, l) : l \in L\}.$$

Then we have:

$$\operatorname{dis}(R) = \sup_{(x,y),(x',y') \in R} |d_{\mathbb{S}^n}(y,y') - d_{\mathbb{S}^0}(x,x')| = |d_{\mathbb{S}^n}(s,s') - d_{\mathbb{S}^0}(p,p)| = \pi.$$

It follows that $d_{\mathrm{GH}}(\mathbb{S}^0, \mathbb{S}^n) \leq \frac{\pi}{2}$. Next we wish to show the reverse inequality. Let T be an arbitrary correspondence between \mathbb{S}^0 and \mathbb{S}^n , and write:

$$P := \{x \in \mathbb{S}^n : (p, x) \in T\}, \qquad Q := \{x \in \mathbb{S}^n : (q, x) \in T\}.$$

Then, $P \neq \emptyset$, $Q \neq \emptyset$, and $\mathbb{S}^n = P \cup Q$, by the definition of correspondences. Now recall the Lusternik-Schnirelmann theorem ([1, p. 117], also see [5, p. 33]): for every family of n+1 closed sets covering \mathbb{S}^n , one of the sets contains a pair of antipodal points. Applying this result by taking n copies of \overline{P} and one copy of \overline{Q} as the cover of \mathbb{S}^n , we get that at least one of the sets \overline{P} and \overline{Q} contains a pair of antipodal points. Without loss of generality, suppose \overline{P} contains a pair of antipodal points (a,a'). Let $(a_n),(a'_n)$ be sequences in P such that for each $n \in \mathbb{N}$, we have $a_n \in B(a,\frac{1}{n})$ and $a'_n \in B(a',\frac{1}{n})$. By the triangle inequality, one has that $|d_{\mathbb{S}^n}(a_n,a'_n)-d_{\mathbb{S}^n}(a,a')| \leq d_{\mathbb{S}^n}(a_n,a)+d_{\mathbb{S}^n}(a'_n,a')<\frac{2}{n}$. Also note that $|d_{\mathbb{S}^n}(a,a')-d_{\mathbb{S}^0}(p,p)|=\pi$. Then we obtain:

$$|d_{\mathbb{S}^n}(a_n,a_n') - d_{\mathbb{S}^0}(p,p)| = |d_{\mathbb{S}^n}(a_n,a_n') - d_{\mathbb{S}^n}(a,a') + d_{\mathbb{S}^n}(a,a') - d_{\mathbb{S}^0}(p,p)| > \pi - \frac{2}{n}.$$

By letting $n \to \infty$, it follows that $\operatorname{dis}(T) \ge \pi$. Since T was an arbitrary correspondence, we obtain $d_{\operatorname{GH}}(\mathbb{S}^0,\mathbb{S}^n) \ge \frac{\pi}{2}$. Thus we obtain $d_{\operatorname{GH}}(\mathbb{S}^0,\mathbb{S}^n) = \frac{\pi}{2}$.

It now follows that the correspondence R defined in the proof of Proposition 1.2 is an optimal correspondence between \mathbb{S}^0 and \mathbb{S}^n . In particular, the definition of R suggests that one may define a geodesic from \mathbb{S}^n to \mathbb{S}^0 by "shrinking" U and L to the north and south poles, respectively.

Proposition 1.3. Let U and L denote the closed upper hemisphere and open lower hemisphere of \mathbb{S}^n , respectively. Also let $\{p,q\}$ denote the two points of \mathbb{S}^0 , and let $\mu \in U$, $\lambda \in L$ denote the north and south poles of \mathbb{S}^n , respectively. For each $t \in (0,1)$, define:

$$U_t := U \cap \overline{B(\mu, (1-t) \cdot \frac{\pi}{2})}, \qquad L_t := L \cap \overline{B(\lambda, (1-t) \cdot \frac{\pi}{2})}$$
$$X_t := U_t \cup L_t, \qquad d_{X_t} := d_{\mathbb{S}^n \mid X_t \times X_t}.$$

Finally define $\gamma:[0,1]\to\mathcal{M}$ by $\gamma(0):=(\mathbb{S}^0,d_{\mathbb{S}^0}),\ \gamma(1):=(\mathbb{S}^n,d_{\mathbb{S}^n}),$ and

$$\gamma(t) := (X_t, d_{X_t}) \text{ for each } t \in (0, 1).$$

Then γ is a geodesic from \mathbb{S}^0 to \mathbb{S}^n .

Proof of Proposition 1.3. First let $t \in (0,1)$, and define a correspondence between \mathbb{S}^0 and X_t by:

$$R_t := \{(p, u) : u \in U_t\} \cup \{(q, l) : l \in L_t\}.$$

Then we have $\operatorname{dis}(R_t) = (1-t)\pi$, and so $d_{\operatorname{GH}}(\mathbb{S}^0, \gamma(t)) \leq \frac{(1-t)\pi}{2}$. Similarly we obtain $d_{\operatorname{GH}}(\gamma(t), \mathbb{S}^n) \leq \frac{t\pi}{2}$. We wish to show that these inequalities are actually equalities. Without loss of generality, suppose that $d_{\mathrm{GH}}(\mathbb{S}^0, \gamma(t)) < \frac{(1-t)\pi}{2}$. Then we obtain:

$$d_{\mathrm{GH}}(\mathbb{S}^0, \mathbb{S}^n) \le d_{\mathrm{GH}}(\mathbb{S}^0, \gamma(t)) + d_{\mathrm{GH}}(\gamma(t), \mathbb{S}^n) < \frac{\pi}{2}.$$

This is a contradiction, by Proposition 1.2. Thus for each $t \in (0,1)$, we have $d_{\mathrm{GH}}(\mathbb{S}^0, \gamma(t)) = \frac{(1-t)\pi}{2}$ and $d_{\mathrm{GH}}(\mathbb{S}^n, \gamma(t)) = \frac{t\pi}{2}$.

Next let $s \in (0,1)$. We wish to show $d_{\mathrm{GH}}(\tilde{\gamma}(s), \gamma(t)) = |t-s| \cdot d_{\mathrm{GH}}(\mathbb{S}^0, \mathbb{S}^n)$. We have two cases: (1) $s \le t$, and (2) s > t. Both cases are similar, so we just show the first case. Before proceeding, notice that since $s \leq t$, we have $U_t \subseteq U_s$ and $L_t \subseteq L_s$, and so $X_t \subseteq X_s$.

We will define some notation for convenience. For each $x \in U_s$ let c_x^{μ} denote the shortest geodesic segment connecting x to the north pole μ . Next, for each $x \in L_s$ let c_x^{λ} denote the shortest geodesic segment connecting x to the south pole λ . Also write $\operatorname{bd}(U_t)$ and $\operatorname{bd}(L_t)$ to denote the boundaries of U_t and L_t . Now define a map $\varphi^U: U_s \to U_t$ by:

$$\varphi^{U}(x) := \begin{cases} x & , x \in U_t \\ c_x^{\mu} \cap \operatorname{bd}(U_t) & , x \in U_s \setminus U_t. \end{cases}$$

Also define a map $\varphi^L: L_s \to L_t$ by:

$$\varphi^{L}(x) := \begin{cases} x & , x \in L_{t} \\ c_{x}^{\lambda} \cap \operatorname{bd}(L_{t}) & , x \in L_{s} \setminus L_{t}. \end{cases}$$

Finally define $\varphi: X_s \to X_t$ as follows:

$$\varphi(x) := \begin{cases} \varphi^{U}(x) & , \ x \in U_s \\ \varphi^{L}(x) & , \ x \in L_s. \end{cases}$$

Observe that for any $x \in X_s$ we have:

$$d_{\mathbb{S}^n}(x,\varphi(x)) \le (1-s) \cdot \frac{\pi}{2} - (1-t) \cdot \frac{\pi}{2} = (t-s) \cdot \frac{\pi}{2}.$$

Now define $T := \{(x, \varphi(x)) : x \in X_s\}$. This is a correspondence between X_s and X_t . By the preceding calculation, we have:

$$\begin{split} \operatorname{dis}(T) &= \sup_{x, x' \in X_s} |d_{\mathbb{S}^n}(x, x') - d_{\mathbb{S}^n}(\varphi(x), \varphi(x'))| \\ &\leq \sup_{x, x' \in X_s} \left(d_{\mathbb{S}^n}(x, \varphi(x)) + d_{\mathbb{S}^n}(x', \varphi(x')) \right) \leq (t - s) \cdot \pi. \end{split}$$

Thus $d_{\mathrm{GH}}(\gamma(s), \gamma(t)) \leq (t-s) \cdot \frac{\pi}{2}$. Similarly, we obtain $d_{\mathrm{GH}}(\gamma(s), \gamma(t)) \leq (s-t) \cdot \frac{\pi}{2}$ when s > t. Thus we obtain:

$$d_{\mathrm{GH}}(\gamma(s), \gamma(t)) \le |t - s| \cdot \frac{\pi}{2} = |t - s| \cdot d_{\mathrm{GH}}(\mathbb{S}^0, \mathbb{S}^n).$$

This inequality must be an equality by Lemma 1.3, which completes the proof that γ is a geodesic.

2. Proof of Theorem 1.2

This section contains our proof showing that $(\mathcal{M}/\sim, d_{\mathrm{GH}})$ is a geodesic space. This fact was established in [6] via an application of Gromov's precompactness theorem. In this paper we present a different, direct proof of this fact based on ideas used by Sturm in the setting of metric measure spaces [9]. In particular, our method of proof provides an explicit construction of geodesics.

Let X and Y be compact metric spaces. Endow $X \times Y$ with the product metric

$$\delta((x,y),(x',y')) := \max(d_X(x,x'),d_Y(y,y')), \text{ for all } (x,y),(x',y') \in X \times Y.$$

Note that $X \times Y$ is compact. Next consider the set of all non-empty closed subsets of $X \times Y$, denoted $\mathcal{C}(X \times Y)$, endowed with the Hausdorff distance δ_{H} arising from δ . It follows from Blaschke's theorem [3] that $\mathcal{C}(X \times Y)$ is also compact.

Lemma 2.1. Let $X \times Y$ be the compact metric space with product metric δ as defined above. Let $R, S \subset X \times Y$ be any two non-empty relations. Then,

- (1) $d_{\mathrm{H}}^{X}(\pi_{1}(R), \pi_{1}(S)) \leq \delta_{\mathrm{H}}(R, S).$ (2) $d_{\mathrm{H}}^{Y}(\pi_{2}(R), \pi_{2}(S)) \leq \delta_{\mathrm{H}}(R, S).$ (3) $|\operatorname{dis}(R) \operatorname{dis}(S)| \leq 4 \delta_{\mathrm{H}}(R, S).$

Here π_1 and π_2 are the natural projections of $X \times Y$ onto X and Y, respectively.

Proof. To show (1), let $\eta > \delta_{H}(R,S)$. Let $x \in \pi_{1}(R)$, and let $y \in Y$ be such that $(x,y) \in R$. Then there exists $(x',y') \in S$ such that $\delta((x,y),(x',y')) < \eta$. Thus $d_X(x,x') < \eta$, where $x' \in \pi_1(S)$. Similarly, given any $u \in \pi_1(S)$, we can find $u' \in \pi_1(R)$ such that $d_X(u, u') < \eta$. Thus $d_H^X(\pi_1(R), \pi_1(S)) < \eta$. Since $\eta > \delta_{\mathrm{H}}(R,S)$ was arbitrary, it follows that $d_{\mathrm{H}}^{X}(\pi_{1}(R),\pi_{1}(S)) \leq \delta_{\mathrm{H}}(R,S)$.

The proof for inequality (2) is similar, so we omit it.

To prove inequality (3), let $\eta > \delta_{\rm H}(R,S)$, and let $\varepsilon \in (\delta_{\rm H}(R,S),\eta)$. Define

$$L := \{(r, s) \in R \times S : \delta(r, s) < \varepsilon\}.$$

Note that, since $\delta_H(R,S) < \varepsilon$, L is a correspondence between R and S. Then,

$$|\operatorname{dis}(R) - \operatorname{dis}(S)|$$

$$= \left| \sup_{(x,y),(x',y') \in R} |d_X(x,x') - d_Y(y,y')| - \sup_{(u,v),(u',v') \in S} |d_X(u,u') - d_Y(v,v')| \right|$$

$$\leq \sup_{\substack{((x,y),(u,v)),\\ ((x',y'),(u',v')) \in L}} \left| |d_X(x,x') - d_Y(y,y')| - |d_X(u,u') - d_Y(v,v')| \right|$$

$$\leq \sup_{\substack{((x,y),(u,v)),\\ ((x',y'),(u',v')) \in L}} \left(|d_X(x,x') - d_X(u,u')| + |d_Y(v,v') - d_Y(y,y')| \right)$$

$$\leq \sup_{\substack{((x,y),(u,v)),\\ ((x',y'),(u',v')) \in L}} \left(d_X(x,u) + d_X(x',u') + d_Y(v,y) + d_Y(v',y') \right) \text{ (Triangle ineq.)}$$

$$\leq \sup_{\substack{((x,y),(u,v)),\\ ((x',y'),(u',v')) \in L}} \left(2\delta((x,y),(u,v)) + 2\delta((x',y'),(u',v')) \right)$$

$$\leq 4\varepsilon < 4\eta.$$

But $\eta > \delta_{H}(R, S)$ was arbitrary. It follows that $|\operatorname{dis}(R) - \operatorname{dis}(S)| \leq 4 \, \delta_{H}(R, S)$. \square

Proof of Proposition 1.1. Let $(\varepsilon_n)_n \downarrow 0$ be an arbitrary sequence in \mathbb{R}_+ . For each n, let X_n, Y_n be ε_n -nets for X and Y respectively. It is a fact that if S is an ε -net in a metric space X, then $d_{GH}(S,X) < \varepsilon$ ([3, Example 7.3.11]). Thus $d_{GH}(X_n,X) \to 0$ and $d_{GH}(Y_n,Y) \to 0$ as $n \to \infty$. Optimal correspondences always exist between finite metric spaces, so for each $n \in \mathbb{N}$, let $R_n \in \mathscr{R}(X_n,Y_n)$ be such that $\operatorname{dis}(R_n) = 2 d_{GH}(X_n,Y_n)$. Since $(R_n)_n$ is a sequence in the compact metric space $\mathcal{C}(X \times Y)$, it contains a convergent subsequence. To avoid double subscripts, we reindex if necessary and let $(R_n)_n$ denote this convergent subsequence. Let $R \in \mathcal{C}(X \times Y)$ denote the δ_H -limit of $(R_n)_n$, i.e. $\lim_{n \to \infty} \delta_H(R_n,R) = 0$. Then by Lemma 2.1,

$$|\operatorname{dis}(R_n) - \operatorname{dis}(R)| \le 4\delta_{\mathrm{H}}(R_n, R) \to 0 \text{ as } n \to \infty.$$

So $dis(R_n) \to dis(R)$. But also,

$$\operatorname{dis}(R_n) = 2 d_{\operatorname{GH}}(X_n, Y_n) \to 2 d_{\operatorname{GH}}(X, Y) \text{ as } n \to \infty,$$

since $|d_{\mathrm{GH}}(X,Y) - d_{\mathrm{GH}}(X_n,Y_n)| \le d_{\mathrm{GH}}(X_n,X) + d_{\mathrm{GH}}(Y_n,Y) \to 0$. Then we have: $\operatorname{dis}(R) = 2 \, d_{\mathrm{GH}}(X,Y)$.

It remains to show that R is a correspondence. Note that for any n,

$$d_{\mathrm{H}}^{X}(X, \pi_{1}(R)) \leq d_{\mathrm{H}}^{X}(X, \pi_{1}(R_{n})) + d_{\mathrm{H}}^{X}(\pi_{1}(R_{n}), \pi_{1}(R))$$

$$\leq d_{\mathrm{H}}^{X}(X, X_{n}) + \delta_{\mathrm{H}}(R_{n}, R) \qquad (\text{By Lemma 2.1}).$$

But the term on the right can be made arbitrarily small, since each X_n is an ε_n net for X and $\lim_{n\to\infty} \delta_{\mathrm{H}}(R_n,R) = 0$. Thus $d_{\mathrm{H}}^X(X,\pi_1(R)) = 0$, and therefore $X = \overline{\pi_1(R)}$. Since R is a closed subset of the compact space $X \times Y$, it is compact, and its continuous image $\pi_1(R)$ is also compact, hence closed (since X is Hausdorff). Thus $\overline{\pi_1(R)} = \pi_1(R) = X$. Similarly, it can be shown that $\pi_2(R) = Y$. Thus R is a correspondence. This concludes the proof.

Proof of Theorem 1.2. Suppose we can find a curve $\gamma : [0,1] \to \mathcal{M}$ such that $\gamma(0) = (X, d_X)$ and $\gamma(1) = (Y, d_Y)$, and for all $s, t \in [0,1]$,

$$d_{\mathrm{GH}}(\gamma(s), \gamma(t)) = |t - s| \cdot d_{\mathrm{GH}}(X, Y).$$

Then we also have $d_{GH}([\gamma(s)], [\gamma(t)]) = |t - s| \cdot d_{GH}([X], [Y])$ for all $s, t \in [0, 1]$, and we will be done. So we will show the existence of such a curve γ .

Let $R \in \mathscr{R}^{\mathrm{opt}}(X,Y)$, i.e. let R be a correspondence between X and Y such that $\mathrm{dis}(R) = 2\,d_{\mathrm{GH}}(X,Y)$. Such a correspondence always exists by Proposition 1.1.

For each $t \in (0,1)$ define $\gamma(t) = (R, d_{\gamma(t)})$ where

$$d_{\gamma(t)}((x,y),(x',y')) = (1-t) \cdot d_X(x,x') + t \cdot d_Y(y,y')$$

for all $(x, y), (x', y') \in R$. Note that for each $t \in (0, 1)$ $d_{\gamma(t)}$ is a legitimate metric on R. We also set $\gamma(0) = (X, d_X)$ and $\gamma(1) = (Y, d_Y)$.

Claim 2.2. For any $s, t \in [0, 1]$,

$$d_{\mathrm{GH}}(\gamma(s), \gamma(t)) \le |t - s| \cdot d_{\mathrm{GH}}(X, Y).$$

Suppose for now that Claim 2.2 holds. Lemma 1.3 implies that, for all $s, t \in [0, 1]$, $d_{\text{GH}}(\gamma(s), \gamma(t)) = |t - s| \cdot d_{\text{GH}}(X, Y)$. Thus it suffices to show Claim 2.2. There are three cases: (i) $s, t \in (0, 1)$, (ii) $s = 0, t \in (0, 1)$, and (iii) $s \in (0, 1), t = 1$. The last two cases are similar, so we just prove (i) and (ii).

For (i), fix $s, t \in (0,1)$. Taking the diagonal correspondence $\blacktriangle \in \mathcal{R}(R,R)$, we get:

$$\begin{split} \operatorname{dis}(\blacktriangle) &= \sup_{(a,a),(b,b) \in \blacktriangle} |d_{\gamma(t)}(a,b) - d_{\gamma(s)}(a,b)| \\ &= \sup_{(x,y),(x',y') \in R} |d_{\gamma(t)}((x,y),(x',y')) - d_{\gamma(s)}((x,y),(x',y'))| \\ &= \sup_{(x,y),(x',y') \in R} |(1-t) \cdot d_X(x,x') + t \cdot d_Y(y,y') \\ &\qquad - (1-s) \cdot d_X(x,x') - s \cdot d_Y(y,y')| \\ &= \sup_{(x,y),(x',y') \in R} |(s-t) \cdot d_X(x,x') - (s-t) \cdot d_Y(y,y')| \\ &= |t-s| \cdot \sup_{(x,y),(x',y') \in R} |d_X(x,x') - d_Y(y,y')| \\ &= 2|t-s| \cdot d_{\mathrm{GH}}(X,Y). \end{split}$$

Finally $d_{\mathrm{GH}}(\gamma(t), \gamma(s)) \leq \frac{1}{2} \mathrm{dis}(\blacktriangle) = |t - s| \cdot d_{\mathrm{GH}}(X, Y)$. This proves case (i) of Claim 2.2.

For (ii), fix $s = 0, t \in (0,1)$. Define $R_X = \{(x,(x,y)) : (x,y) \in R\}$. Then R_X is a correspondence in $\mathcal{R}(X,R)$.

$$dis(R_X) = \sup_{(x,(x,y)),(x',(x',y')) \in R_X} |d_X(x,x') - (1-t) \cdot d_X(x,x') - t \cdot d_Y(y,y')|$$

$$= \sup_{(x,(x,y)),(x',(x',y')) \in R_X} |d_X(x,x') - d_Y(y,y')| \cdot t$$

$$= t \cdot dis(R) = 2t \cdot d_{GH}(X,Y).$$

Thus $d_{\mathrm{GH}}(X, \gamma(t)) \leq t \cdot d_{\mathrm{GH}}(X, Y)$. The proof for case (iii), i.e. that $d_{\mathrm{GH}}(\gamma(s), Y) \leq |1-s| \cdot d_{\mathrm{GH}}(X, Y)$, is similar. Thus Claim 2.2 follows. The theorem now follows. \square

3. Discussion

While we provide an explicit construction of straight-line geodesics, it is natural to ask the following: can we characterize other classes of geodesics in $(\mathcal{M}/\sim, d_{\rm GH})$? In Section 1.1.1, we constructed infinite families of deviant (i.e. non-unique) geodesics between Δ_1 and Δ_n . In Section 1.1.2, we provided a parametric construction by which the straight-line geodesic between Δ_1 and Δ_n could be made to branch off into arbitrarily many nodes at arbitrarily many locations.

As stated at the end of Section 1.1, the existence of branching and deviant geodesics shows the negative result that $(\mathcal{M}/\sim, d_{\rm GH})$ cannot have curvature bounded from above or below. In light of this result, it is interesting to point out the work of Sturm showing that the space of metric measure spaces [9] has nonnegative curvature when equipped with an L^2 -Gromov-Wasserstein metric.

4. Acknowledgments

We thank Prof. Vladimir Zolotov for pointing out the existence of branching geodesics to us. We also thank the referees for their helpful comments.

References

- Béla Bollobás, The art of mathematics: Coffee time in Memphis, Cambridge University Press, 2006.
- 2. Martin R Bridson and André Haefliger, Metric spaces of non-positive curvature, vol. 319, Springer Science & Business Media, 2011.
- 3. Dmitri Burago, Yuri Burago, and Sergei Ivanov, A course in metric geometry, AMS Graduate Studies in Math., vol. 33, American Mathematical Society, 2001.
- 4. Misha Gromov, Metric structures for Riemannian and non-Riemannian spaces, Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., Boston, MA, 1999.
- 5. Allen Hatcher, Algebraic topology, Cambridge University Press, 2002.
- Alexandr Ivanov, Nadezhda Nikolaeva, and Alexey Tuzhilin, The Gromov-Hausdorff metric on the space of compact metric spaces is strictly intrinsic, arXiv preprint arXiv:1504.03830 (2015).
- 7. Vladimir Pestov, Dynamics of infinite-dimensional groups: the ramsey-dvoretzky-milman phenomenon, vol. 40, American Mathematical Soc., 2006.
- 8. Peter Petersen, Riemannian geometry, vol. 171, Springer Science & Business Media, 2006.
- 9. Karl-Theodor Sturm, The space of spaces: curvature bounds and gradient flows on the space of metric measure spaces, arXiv preprint arXiv:1208.0434 (2012).

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 100 MATH TOWER, 231 WEST 18TH AVENUE, COLUMBUS, OH 43210. PHONE: (614) 292-4975, FAX: (614) 292-1479 E-mail address: chowdhury.57@osu.edu

Department of Mathematics, The Ohio State University, 100 Math Tower, 231 West 18th Avenue, Columbus, OH 43210. Phone: (614) 292-4975, Fax: (614) 292-1479 E-mail address: memoli@math.osu.edu