A metric on the space of weighted graphs

Hamed Daneshpajouh Hamid Reza Daneshpajouh * Farzad Didehvar †

June 16, 2009

Abstract

In this paper we offer a metric similar to graph edit distance which measures the distance between two (possibly infinite)weighted graphs with finite norm (we define the norm of a graph as the sum of absolute values of its edges). The main result is the completeness of the space. Some other analytical properties of this space are also investigated.

1 Introduction

Many objects can be demonstrated with weighted graphs. In any collection of objects of similar nature a way to quantify the difference between objects may be desired (For instance if we were to select the most similar objects to a given object from a database). In the theoretical side one common way is to develop a metric on the space of objects in demand. One way to build a metric, is to define some operations that transform the members of the space to one another, and assign a cost to each operation then define the distance between two objects to be the minimum cost that must be payed to transform the first object to the second via a sequence of the defined operations. Such metrics sometimes are referred to as "Edit distance". Two examples of them are the "Levenshtein edit distance" [1] on strings and "Graph edit distance" [2] on the space of finite graphs. This paper extends the Graph edit distance to the space of "countable weighted graphs with finite norm" and investigates some topological properties of the space.

2 Priliminaries and intuitive examples

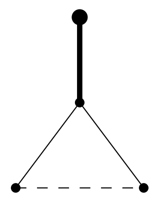
In this chapter, we introduce the concepts intuitively. The main question here is "given two graphs, how much they differ?". Based on this question we could define different distances, we choose here "Graph edit distance" and we generalize it to infinite graphs.

Example 1. *Consider the following two graphs*

*Department of Mathematics and Computer Science, University of Tehran. Email: h.r.daneshpajouh@khayam.ut.ac.ir.

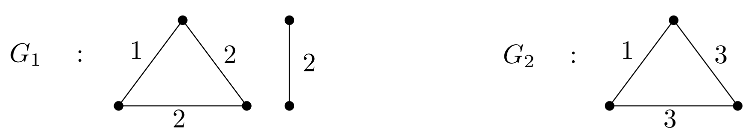
†Department of Mathematics and Computer science, Amirkabir University of Technology. Email: didehvar@aut.ac.ir, fr@ipm.ir.

One can transform G_1 to G_2, by adding a vertex and an edge to G_1 and deleting an edge from it

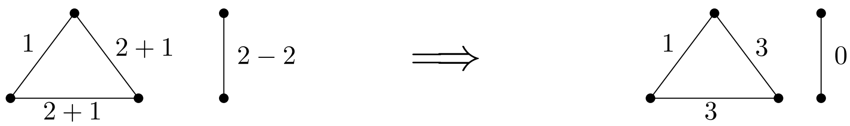


Given two graphs, it is possible to transform one to the other by addition and deletion of some vertices and edges. The minimum number of edge addition and deletions in such a process is the distance between the two graphs and is denoted by $\mathrm{d}(G_1, G_2)$. In the above example $\mathrm{d}(G_1, G_2) = 2$, because we added an edge and deleted one. It is clear that if two graphs differ only in isolated vertices, then by this definition their distance is zero.

Example 2. *Consider the following weighted graphs*

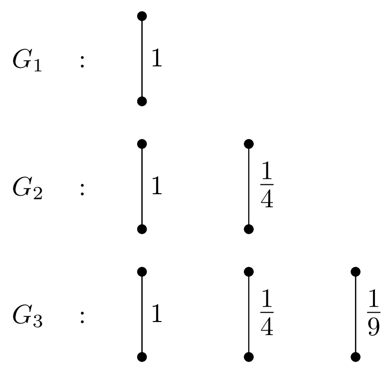


We transform G_1 to G_2 as follows



The right-hand graph, after removing the zero weighted edge and it's endpoints is same as G_2. We define the distance between two edge-weighted graphs to be the minimum amount of edge-weight modifications required to transform one to the other (zero-weighted edges and isolated vertices could be added and deleted for free) , in this example $\mathrm{d}(G_1, G_2) = |1| + |1| + |-2| = 4$.

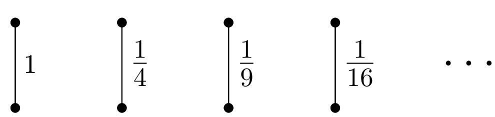
Example 3. *We construct the sequence $\{G_n\}$ as follows*



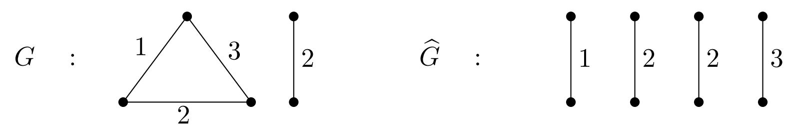
If $n < m$ then clearly

$$\mathrm{d}(G_n, G_m) \leq \sum_{t=n+1}^{m} \frac{1}{t^2}$$

hence $\{G_n\}$ is a Cauchy sequence. On the other hand $\{G_n\}$ does not approach to a finite graph (to prove that let G be a graph with m edges, and show that for each $n > m$, $\mathrm{d}(G_n, G_m) \geq 1/(m+1)$), implying that the space of finite weighted graphs is not complete. But this sequence approaches to the following infinite graph



Example 4. *Consider the following two graphs*



$\widehat{G}$ *is obtained from* G *by removing the joints. You should be able to verify the following inequality for arbitrary graphs intuitively.*

$$\mathrm{d}(\widehat{G}, \widehat{H}) \leq \mathrm{d}(G, H)$$

3 Developing the metric mathematically

We first define the distance between two labeled graphs and then define unlabeled graphs as the equivalence classes of labeled graphs. The main result in this chapter is to show that the introduced distance provides us a metric space.

Definition 1. *Fix* $V = \{v_1, v_2, v_3, \cdots\}$ *as the vertex set. We show the set of all 2-element subsets of* V *by* E *(edge set). A labeled graph is a function* $w : E \to \mathbb{R}$*. The value* $w(e)$*, is the weight of* e *(in* w*). A nonzero edge (of* w*), is one which it's weight is nonzero. An isolated vertex of* w *is a vertex that has no nonzero edge attached.* w *is standard if it has infinitely many isolated vertices. We denote by* W *the set of all labeled graphs, and by* W^s *the set of all standard labeled graphs. For* $w, w' \in W$*, the norm of* w*, the distance between* w *and* w' *and the sets* W_0 *and* W_0^s *are defined as follows*

$$|w| = \textstyle\sum_{e \in E} |w(e)|$$

$$\mathrm{d}(w, w') = |w - w'|$$

$$W_0 = \Big\{ w \in W \mid |w| < \infty \Big\}$$

$$W_0^s = W_0 \cap W^s$$

W *is a vector space and* W_0 *is a normed subspace of it which, under metric* d *is isometric to* $l_1(\mathbb{R})$*.* $w, w' \in W$ *are isomorphic(*$w \sim w'$*), when there exists a bijection* $f : V \to V$ *such that for every edge* $uv \in E$*,* $w(uv) = w'(f(u), f(v))$*.*

We use standard graphs in constructing unlabeled weighted graphs, it has several benefits, in particular with an infinite number of isolated vertices we don't need to delete or add a vertex.

Definition 2. *Let* $\mathcal{G} = W^s / \sim$ *and* $\mathcal{G}_0 = W_0^s / \sim$*. The elements of* $\mathcal{G}$ *are named unlabeled graphs. For* $w \in W$*, we show the equivalence class of* w *by* $\widetilde{w}$ *and define*

$$|\widetilde{w}| = |w|$$

When no confusion can arise we use the term “graph” instead of “labeled graph” and “unlabeled graph”. A graph with all edge weights equal to 0, simply is denoted by 0. We sometimes use common graph theory notions here, converting them to match our definitions is not difficult, for example “to delete an edge” means “to change its weight to zero”.

Definition 3. *We denote by $S(V)$ the set of all bijections on V. Suppose $\sigma \in S(V)$ and $e = uv \in E$, let us define $\sigma(e) = \{\sigma(u), \sigma(v)\}$. Also for $w \in W$, define $w^\sigma \in W$ as follows*

$$\forall e \in E \quad w^\sigma(\sigma(e)) = w(e)$$

or equivalently

$$\forall e \in E \quad w^\sigma(e) = w(\sigma^{-1}(e))$$

obviously for $w, w' \in W$

$$w \sim w' \iff \exists \sigma \in S(V) : w' = w^\sigma$$

$$\widetilde{w} = \Big\{ w^\sigma \mid \sigma \in S(V) \Big\}$$

Lemma 1. *Given $w, w' \in W$ and $\sigma, \gamma \in S(V)$, the following equations hold*

$$\mathrm{d}(w^\sigma, w'^\sigma) = \mathrm{d}(w, w')$$

$$(w^\sigma)^\gamma = w^{\gamma \circ \sigma}$$

Definition 4. *Given two graphs $G, G' \in \mathcal{G}$, we define their distance as follows*

$$\mathrm{d}(G, G') = \inf \Big\{ \mathrm{d}(w, w') \mid w \in G, w' \in G' \Big\}$$

Also for $w \in G$ we define

$$\mathrm{d}(G', w) = \mathrm{d}(w, G') = \inf \Big\{ \mathrm{d}(w, w') \mid w' \in G' \Big\}$$

Lemma 2. *For $w_1, w_2 \in W^s$*

$$\mathrm{d}(\widetilde{w}_1, \widetilde{w}_2) = \mathrm{d}(w_1, \widetilde{w}_2) = \mathrm{d}(\widetilde{w}_1, w_2)$$

proof.

$$\mathrm{d}(w_1^{\sigma_1}, w_2^{\sigma_2}) = \mathrm{d}(w_1, w_2^{\sigma_1^{-1} \circ \sigma_2}) = \mathrm{d}(w_1^{\sigma_2^{-1} \circ \sigma_1}, w_2) \Longrightarrow$$

$$\Big\{ \mathrm{d}(w_1^{\sigma_1}, w_2^{\sigma_2}) \mid \sigma_1, \sigma_2 \in S(V) \Big\} = \Big\{ \mathrm{d}(w_1, w_2^{\alpha}) \mid \alpha \in S(V)) \Big\} =$$

$$\Big\{ \mathrm{d}(w_1^{\beta}, w_2) \mid \beta \in S(V) \Big\}$$

■

Lemma 3. *For $w_1, w_2, w_3 \in W^s$*

$$\mathrm{d}(\widetilde{w}_1, \widetilde{w}_3) \leq \mathrm{d}(\widetilde{w}_1, \widetilde{w}_2) + \mathrm{d}(\widetilde{w}_2, \widetilde{w}_3)$$

proof.

$$\mathrm{d}(\widetilde{w}_1, \widetilde{w}_2) + \mathrm{d}(\widetilde{w}_2, \widetilde{w}_3) = \mathrm{d}(\widetilde{w}_1, w_2) + \mathrm{d}(w_2, \widetilde{w}_3) \geq \mathrm{d}(\widetilde{w}_1, \widetilde{w}_3)$$

■

Definition 5. *Let $w \in W$, $A \subseteq \mathbb{R}$ and $E' \subseteq E$. We define graphs $\mathrm{Cut}(w, A), \mathrm{Cut}(w, E') \in W$ as follows*

$$\mathrm{Cut}(w, A)(e) = \begin{cases} w(e) & w(e) \in A \\ 0 & \text{otherwise} \end{cases}$$

this means to delete all edges with weights outside of A

$$\mathrm{Cut}(w, E')(e) = \begin{cases} w(e) & e \in E' \\ 0 & \text{otherwise} \end{cases}$$

also for $\epsilon \geq 0$

$$\mathrm{Cut}(w, \epsilon) = \mathrm{Cut}(w, (-\infty, -\epsilon] \cup [\epsilon, \infty))$$

moreover for $w \in W^s$ we have two more definitions

$$\mathrm{Cut}(\widetilde{w}, A) = \widetilde{\mathrm{Cut}(w, A)} \qquad \textit{and} \qquad \mathrm{Cut}(\widetilde{w}, \epsilon) = \widetilde{\mathrm{Cut}(w, \epsilon)}$$

Definition 6. *The spectrum of a graph is the set of all of it's edge weights*

$$\mathrm{Spec}(w) = \mathrm{Spec}(\widetilde{w}) = \Big\{ w(e) \mid e \in E \Big\}$$

Lemma 4. *If $w \in W_0^s$ then $\mathrm{Spec}(w)$ is a countable and compact set, furthermore the only possible limit point of it is 0.*

Lemma 5. *If $G, H \in \mathcal{G}_0$ and $\mathrm{d}(G, H) = 0$ then for every $\epsilon \geq 0$, $\mathrm{Cut}(G, \epsilon) = \mathrm{Cut}(H, \epsilon)$.*

proof. Let

$$A = \Big\{ |x - y| \mid x \in \mathrm{Spec}(G), y \in \mathrm{Spec}(H), x \neq y, (|x| \geq \epsilon \text{ or } |y| \geq \epsilon) \Big\}$$

If $A = \emptyset$ then $\mathrm{Cut}(G, \epsilon) = 0 = \mathrm{Cut}(H, \epsilon)$, otherwise we can define

$$\delta = \min A$$

Clearly $\delta > 0$. Choose $w \in G$ and $w' \in H$ such that $|w - w'| < \delta$. For $e \in E$, $|w(e) - w'(e)| < \delta$, and two cases are possible
case i: $|w(e)|, |w'(e)| < \epsilon$, which implies

$$\mathrm{Cut}(w, \epsilon)(e) = 0 = \mathrm{Cut}(w', \epsilon)(e)$$

case ii: $w(e) \geq \epsilon$ or $w'(e) \geq \epsilon$, in this case $w(e) = w'(e)$ because otherwise $|w(e) - w'(e)| \in A$, and hence $\delta < \delta$ which is a contradiction. Thus

$$\mathrm{Cut}(w, \epsilon)(e) = w(e) = w'(e) = \mathrm{Cut}(w', \epsilon)(e)$$

therefore in each case, $\mathrm{Cut}(w, \epsilon) = \mathrm{Cut}(w', \epsilon)$ and consequently $\mathrm{Cut}(G, \epsilon) = \mathrm{Cut}(H, \epsilon)$. ■

In the following lemma which is known as Konig infinity lemma, please forget our notion of a graph, just take it as in ordinary graph theory texts.

Lemma 6. *$A_1, A_2, A_3, \cdots$ are nonempty, finite and disjoint sets, and G is a graph with $\bigcup_{n=1}^{\infty} A_n$ as vertex set, such that for every n, every vertex in A_{n+1} has a neighbour in A_n. G contains a ray $a_1 a_2 a_3 \cdots$ with $a_n \in A_n$. (A ray is a sequence of different vertices each of which adjacent to it's successor)*

Theorem 1. *if $w, w' \in W_0^s$ and $\mathrm{d}(\widetilde{w}, \widetilde{w}') = 0$ then $\widetilde{w} = \widetilde{w}'$, and consequently d is a metric on $\mathcal{G}_0$*

proof. Assume that $w_n = \mathrm{Cut}(w, \frac{1}{n})$ and $w'_n = \mathrm{Cut}(w', \frac{1}{n})$. The above lemma implies that $w_n \sim w'_n$. Denote by U_n and U'_n the sets of non-isolated vertices of w_n and w'_n, and by A_n the set of all pairs (n, f) in which f is an isomorphism between nonzero parts of w_n and w'_n

$$A_n = \{(n, f) | f : U_n \leftrightarrow U'_n, \forall u, v \in U_n (u \neq v \Rightarrow w_n(uv) = w'_n(f(u), f(v)))\}$$

Since $w_n \sim w'_n$ and U_n, U'_n are finite, A_n is nonempty and finite. Define a graph with vertex set $\bigcup_{n=1}^{\infty} A_n$ and edge set $\Big\{\{(n, f), (n+1, g)\} \mid f \subseteq g\Big\}$. Consider $(n+1, g) \in A_{n+1}$. Let f be the restriction of g to U_n, it is easily seen that $(n, f) \in A_n$ and (n, f) is a neighbor of $(n+1, g)$ so each vertex in A_{n+1} has a neighbour in A_n. Then according to Konig infinity lemma, there is an infinite sequence $(1, f_1), (2, f_2), (3, f_3), \cdots$ of vertices such that for each n, the vertex (n, f_n) is adjacent to the vertex $(n+1, f_{n+1})$, i.e $f_1 \subseteq f_2 \subseteq f_3 \subseteq \cdots$. We put $f = \bigcup_{n=1}^{\infty} f_n$. f is an isomorphism between nonzero parts of w and w'. Since both w and w' have a countable number of isolated vertices, f can be extended to an isomorphism between w and w'. ■

4 Completeness of $\mathcal{G}_0$

In this chapter and the next one we try to find some topological properties of $\mathcal{G}_0$. The main result in this chapter is the completeness of $\mathcal{G}_0$.

Theorem 2. *Let $G, G_1, G_2, \cdots \in \mathcal{G}_0$. The following are equivalent*

1. *$G_n \to G$*

2. *for every $w \in G$ there is a sequence $w_n \in G_n$ such that $w_n \to w$*

3. *there is a $w \in G$ and a sequence $w_n \in G_n$ such that $w_n \to w$*

proof. (1) $\Rightarrow$ (2): Let $w \in G$, since

$$\mathrm{d}(G_n, G) = \mathrm{d}(G_n, w) = \inf \Big\{ \mathrm{d}(w', w) \mid w' \in G_n \Big\}$$

the elements $w_n \in G_n$ exist such that $\mathrm{d}(w_n, w) \leq \mathrm{d}(G_n, G) + \frac{1}{n}$, therefore $w_n \to w$. (2) $\Rightarrow$ (3) is evident
(3) $\Rightarrow$ (1) It follows from the inequality

$$\mathrm{d}(G_n, G) \leq \mathrm{d}(w_n, w)$$

■

Theorem 3. *If the sequence $\{w_n\} \subseteq W_0^s$ be convergent to a graph in W_0, then the sequence $\{\widetilde{w}_n\}$ is convergent in $\mathcal{G}_0$.*

proof. Suppose $w_n \to w$, we shall define the graphs w'_n and w' by means of the following equations

$$w'_n(v_i v_j) = \begin{cases} w_n(v_{i/2} v_{j/2}) & i, j \text{ are even} \\ 0 & \text{otherwise} \end{cases}$$

$$w'(v_i v_j) = \begin{cases} w(v_{i/2} v_{j/2}) & i, j \text{ are even} \\ 0 & \text{otherwise} \end{cases}$$

we simply observe that w' is a standard graph and $w_n \sim w'_n$ and $\mathrm{d}(w'_n, w') = \mathrm{d}(w_n, w)$ so $w'_n \to w'$, therefore $\widetilde{w}'_n \to \widetilde{w}'$ and accordingly $\widetilde{w}_n \to \widetilde{w}'$. ■

Definition 7. *Suppose that $w, w' \in W$. w is a subgraph of w' ($w \preceq w'$) when*

$$\forall e \in E \quad \Big(w(e) = 0 \quad \vee \quad w(e) = w'(e) \Big)$$

also we say, G is a subgraph of G' ($G \preceq G'$) where $G, G' \in \mathcal{G}$, if one of these equivalent statements holds

$$\exists w \in G \; \exists w' \in G' \quad \Big(w \preceq w' \Big)$$

$$\forall w \in G \; \exists w' \in G' \quad \Big(w \preceq w' \Big)$$

$$\forall w' \in G' \; \exists w \in G \quad \Big(w \preceq w' \Big)$$

Clearly if $w, w' \in W_0$ and $w \prec w'$ then $|w| < |w'|$. Also $\preceq$ is a partial order on W.

Theorem 4. *$\preceq$ is a partial order on $\mathcal{G}_0$.*

proof. The transitive and reflexive properties are consequences of the similar properties in W. The proof of the antisymmetric property: If this property fails, then $G \prec G' \prec G$ holds for some $G, G' \in \mathcal{G}_0$, implying $|G| < |G'| < |G|$, which is impossible. ■

Theorem 5. *If $G, G' \in \mathcal{G}_0$, then*

1. $\mathrm{d}(G, G') \geq ||G| - |G'||$

2. *If* $G \preceq G'$ *then* $\mathrm{d}(G, G') = |G'| - |G|$

proof. 1) For each $w \in G$ and $w' \in G'$ we have

$$|w - w'| \geq ||w| - |w'|| = ||G| - |G'||$$

so $\mathrm{d}(G, G') \geq ||G| - |G'||$.
2) Choose $w \in G$ and $w' \in G'$ such that $w \preceq w'$, clearly

$$|w' - w| = |w'| - |w| = |G'| - |G|$$

which in combination with (1) gives the result. ■

If $G_1 \preceq G_2 \preceq G_3 \preceq \cdots$ is an increasing sequence in $\mathcal{G}$, then there exists an increasing sequence $w_1 \preceq w_2 \preceq w_3 \preceq \cdots$ such that for every $n \in N$, $w_n \in G_n$. It is enough to select w_1 from G_1 then construct other terms inductively.

Theorem 6. *Suppose that* $w, w_1, w_2, w_3, \cdots \in W_0$, $w_n \to w$ *and* $\{a_n\}$ *is a sequence of nonnegative real numbers converging to zero, it follows that* $\mathrm{Cut}(w_n, a_n) \to w$.

proof. We take $\epsilon > 0$ and select $\delta > 0$ satisfying $\mathrm{Cut}(w, [-\delta, \delta]) < \epsilon$ and choose $N_1 \in N$ such that

$$\forall n \geq N_1 \quad a_n < \delta$$

we also set $A = \Big\{ e \in E \mid |w(e)| > \delta \Big\}$. Since A is finite, we can select N_2 and N_3 in such a way that

$$\forall n \geq N_2 \; \forall e \in A \quad |w_n(e)| > \delta$$

$$\forall n \geq N_3 \quad |w_n - w| < \epsilon$$

Now for $n \geq \max\{N_1, N_2, N_3\}$

$$|\mathrm{Cut}(w_n, a_n)| \geq |\mathrm{Cut}(w_n, A)| \geq |\mathrm{Cut}(w, A)| - |\mathrm{Cut}(w - w_n, A)|$$

$$\geq |w| - |\mathrm{Cut}(w, [-\delta, \delta])| - |w - w_n| \geq |w| - \epsilon - \epsilon = |w| - 2\epsilon$$

$$\Longrightarrow$$

$$|w - \mathrm{Cut}(w_n, a_n)| \leq |w - w_n| + |w_n - \mathrm{Cut}(w_n, a_n)|$$

$$\leq \epsilon + |w_n| - |\mathrm{Cut}(w_n, a_n)| \leq \epsilon + |w| + \epsilon - (|w| - 2\epsilon) = 4\epsilon$$

■

Theorem 7. *In* W_0 *(or* $\mathcal{G}_0$*), any increasing bounded sequence(with respect to* $\preceq$*) is convergent.*

proof. Suppose that $\{w_n\}$ is an increasing bounded sequence in W_0. We define the graph w as follows

$$w(e) = \lim_{n\to\infty} w_n(e)$$

The sequence $w_n(e)$ is ultimately constant, so the limit exists. It is easily seen that

$$|w| = \lim_{n\to\infty} |w_n| < \infty \quad \Longrightarrow \quad w \in W_0$$

Given the fact that $w_n \preceq w$ for any n, we have

$$\lim_{n\to\infty} \mathrm{d}(w, w_n) = \lim_{n\to\infty} (|w| - |w_n|) = 0$$

therefore $w_n \to w$.
Now, suppose $\{G_n\}$ is an increasing bounded sequence in $\mathcal{G}_0$. Corresponding to this sequence, there is an increasing sequence $\{w_n\} \subseteq W_0$ such that $w_n \in G_n$. The convergence of $\{G_n\}$ is a result of the convergence of $\{w_n\}$. ■

Definition 8. *A graph in which no two nonzero edges are adjacent is called a jointless one. We denote by $\widehat{\mathcal{G}}$ and $\widehat{\mathcal{G}}_0$ the sets of jointless graphs in $\mathcal{G}$ and $\mathcal{G}_0$. Given a graph $G \in \mathcal{G}_0$, there is a unique member of $\widehat{\mathcal{G}}_0$ which has the same edge weights as G (with same multiplicity), we denote it by $\widehat{G}$, see example 4.*

Definition 9. *Suppose $G \in \widehat{\mathcal{G}}$ and $a \in \mathrm{Spec}(G)$. The unique graph obtained from G by deleting an edge with weight a is denoted by $\mathrm{Del}(G, a)$.*

Definition 10. *Suppose $\widetilde{w} = G \in \mathcal{G}_0$. We define $\mathrm{Code}(w) = \mathrm{Code}(G) = f$ where $f \in l_1(\mathbb{R})$ is constructed by induction as follows: set $G_0 = G$ and define*

$$f(n) = \begin{cases} \max\ \mathrm{Spec}(G_{n-1}) & n \text{ is odd} \\ \min\ \mathrm{Spec}(G_{n-1}) & n \text{ is even} \end{cases} \qquad G_n = \mathrm{Del}(G_{n-1}, f(n))$$

Example 5. *Suppose that G is the jointless graph that has an edge with weight $\frac{1}{2^n}$ for each $n \geq 1$ and two edges of weight -1 , then*

$$\mathrm{Code}(G) = (\frac{1}{2}, -1, \frac{1}{4}, -1, \frac{1}{8}, 0, \frac{1}{16}, 0, \cdots)$$

Lemma 7. *Suppose $G, H \in \widehat{\mathcal{G}}_0$, $x_1 = \max \mathrm{Spec}(G)$ and $x_2 = \max \mathrm{Spec}(H)$ (or $x_1 = \min \mathrm{Spec}(G)$ and $x_2 = \min \mathrm{Spec}(H)$) and $G' = \mathrm{Del}(G, x_1)$ and $H' = \mathrm{Del}(H, x_2)$, then*

$$\mathrm{d}(G, H) = \mathrm{d}(G', H') + |x_1 - x_2|$$

proof. It is clear that $\mathrm{d}(G,H) \le \mathrm{d}(G', H') + |x_1 - x_2|$. To show the inverse, we take $w_1 \in G$ and $w_2 \in H$. Let $x_1 = w_1(e_1)$, $x_2 = w_2(e_2)$, $y_1 = w_1(e_2)$ and $y_2 = w_2(e_1)$ and define the graph w as follows

$$w(e) = \begin{cases} x_1 & e = e_2 \\ y_1 & e = e_1 \\ w_1(e) & \text{otherwise} \end{cases}$$

according to the assumption, $y_1 \le x_1$ and $y_2 \le x_2$ (or $x_1 \le y_1$ and $x_2 \le y_2$). Using these relations one can easily show that

$$\mathrm{d}(w_1, w_2) - \mathrm{d}(w, w_2) = |x_1 - y_2| + |x_2 - y_1| - |x_1 - x_2| - |y_2 - y_1| \ge 0$$

$$\implies \quad \mathrm{d}(w_1, w_2) \ge \mathrm{d}(w, w_2) \ge |x_1 - x_2| + \mathrm{d}(G', H')$$

■

Theorem 8. *For every $G, H \in \mathcal{G}_0$ we have*

$$\mathrm{d}(G,H) = |\mathrm{Code}(G) - \mathrm{Code}(H)|$$

proof. Let G_n be the sequence related to graph G in definition 10, and relate a similar sequence H_n to H. We set $\mathrm{Code}(G) = g$ and $\mathrm{Code}(H) = h$. Applying the preceding lemma n times, we obtain

$$\mathrm{d}(G,H) = \sum_{i=1}^{n} |g(n) - h(n)| + \mathrm{d}(G_n, H_n)$$

Since the sequences $|G_n|$ and $|H_n|$ are convergent to 0, $\mathrm{d}(G_n, H_n) \to 0$ and consequently

$$\mathrm{d}(G,H) = \sum_{i=1}^{\infty} |g(n) - h(n)| = |\mathrm{Code}(G) - \mathrm{Code}(H)|$$

■

Lemma 8. *$\widehat{\mathcal{G}_0}$ is a complete subspace of $\mathcal{G}_0$*

proof. Set $B = \left\{ \mathrm{Code}(G) \mid G \in \widehat{\mathcal{G}_0} \right\}$ it is easily seen that

$$B = \Big\{ f \in l_1 \mid \forall n \ \Big(f(2n-1) \ge 0, f(2n) \le 0,$$

$$f(2n-1) \ge f(2n+1), f(2n) \le f(2n+2) \Big) \Big\}$$

and B is a closed subset and consequently a complete subset of $l_1(\mathbb{R})$. According to the theorem 8, $\mathrm{Code} : \widehat{\mathcal{G}_0} \to B$ is an onto isometry, so $\widehat{\mathcal{G}_0}$ is also complete. ■

By a correspondance between two graphs $G_1, G_2 \in \mathcal{G}$ we mean a choice of two members $w_1 \in G_1$ and $w_2 \in G_2$.

Theorem 9. *$\mathcal{G}_0$ is a complete metric space*

proof. Suppose that G_n is a Cauchy sequence. Since $\mathrm{d}(\widehat{G}_n, \widehat{G}_m) \leq \mathrm{d}(G_n, G_m)$, so the sequence $\{\widehat{G}_n\}$ is also Cauchy. Therefore according to the preceding lemma, the sequence $\{\widehat{G}_n\}$ is convergent to a jointless graph F. Set $A = \mathrm{Spec}(F)$ and let $w_n \in G_n$. Suppose w'_n is obtained from w_n by rounding the weight of each edge to the closest number in A (if the weight of an edge has the least difference with two numbers in A, we choose one of them arbitrarily). Set $H_n = \widetilde{w}'_n$, we claim that $\mathrm{d}(H_n, G_n) \to 0$. In fact, it is easily seen that $\mathrm{d}(G_n, H_n) \leq \mathrm{d}(w_n, w'_n) \leq \mathrm{d}(\widehat{G}_n, F)$, which proves the claim. Therefore, it is enough to show $\{H_n\}$'s convergence instead of $\{G_n\}$'s. First let us show that for all $\epsilon > 0$ there exists a natural number M such that

$$\forall m \geq M \quad \mathrm{Cut}(H_m, \epsilon) = \mathrm{Cut}(H_M, \epsilon) \tag{1}$$

To prove that we set $A_\epsilon = \Big\{x \in A \mid \ |x| \geq \epsilon\Big\}$ and $\delta = \min\Big\{\mathrm{d}(x, A\backslash\{x\}) \mid x \in A_\epsilon\Big\}$. Since A does not have a nonzero limit point and A_ϵ is finite, so $\delta > 0$. The relation $\mathrm{d}(H_n, G_n) \to 0$ shows that $\{H_n\}$ is a Cauchy sequence so there is a $M \in \mathbb{N}$ such that $\mathrm{d}(H_M, H_m) < \delta$ for every $m \geq M$. Now, if $\mathrm{Cut}(H_m, \epsilon) \neq \mathrm{Cut}(H_M, \epsilon)$ for one $m \geq M$, in every correspondence between H_m and H_M we get an edge which has two different weights in the two graphs, one of which from A_ϵ, and the other from A, and hence $\mathrm{d}(H_M, H_m) \geq \delta$, which is a contradiction. Choose a strictly increasing sequence $\{M_n\}$ such that for each n, $\epsilon = \frac{1}{n}$ and $M = M_n$ satisfy the equation (1). It is evident that for each n, $\mathrm{Cut}(H_{M_n}, \frac{1}{n}) \preceq \mathrm{Cut}(H_{M_{n+1}}, \frac{1}{n+1})$. Then, since $\Big|\mathrm{Cut}(H_{M_n}, \frac{1}{n})\Big| \leq |H_{M_n}|$, so the sequence $\{\mathrm{Cut}(H_{M_n}, \frac{1}{n})\}$ is bounded and consequently converges to a graph G. We have

$$\begin{gathered}
\mathrm{d}(\widehat{H}_n, \widehat{G}_n) \leq \mathrm{d}(H_n, G_n) \Longrightarrow \mathrm{d}(\widehat{H}_n, \widehat{G}_n) \to 0 \\
\Longrightarrow \widehat{H}_n \to F \Longrightarrow \widehat{H}_{M_n} \to F \Longrightarrow \\
\mathrm{Cut}(\widehat{H}_{M_n}, \tfrac{1}{n}) \to F \Longrightarrow \mathrm{d}(\widehat{H}_{M_n}, \mathrm{Cut}(\widehat{H}_{M_n}, \tfrac{1}{n})) \to 0
\end{gathered} \tag{2}$$

Also

$$\begin{gathered}
\mathrm{d}(H_{M_n}, \mathrm{Cut}(H_{M_n}, \tfrac{1}{n})) = |H_{M_n}| - \Big|\mathrm{Cut}(H_{M_n}, \tfrac{1}{n})\Big| \\
= \Big|\widehat{H}_{M_n}\Big| - \Big|\mathrm{Cut}(\widehat{H}_{M_n}, \tfrac{1}{n})\Big| = \mathrm{d}(\widehat{H}_{M_n}, \mathrm{Cut}(\widehat{H}_{M_n}, \tfrac{1}{n}))
\end{gathered} \tag{3}$$

(2) and (3) conclude that

$$\mathrm{d}(H_{M_n}, \mathrm{Cut}(H_{M_n}, \frac{1}{n})) \to 0$$

Therefore $H_{M_n} \to G$, so H_n has a convergent subsequence and consequently, it is convergent itself. ■

5 Examining the space for some other common topological properties

Besides the completeness of $\mathcal{G}_0$ there are some other important properties of the space, we discuss a few of them here. Note that, in this chapter by a finite graph we mean one that has

only finitely many non-isolated vertices. Also in a metric space $\mathcal{M}$ we denote the ball with center x and radius r by $B_{\mathcal{M}}(x, r)$.

Theorem 10. *$\mathcal{G}_0$ is separable.*

proof. The set of all finite graphs with rational edge weights is a dense subset of $\mathcal{G}_0$. ■ By a finite graph we mean one that has only finitely many non-isolated vertices. Also in a metric space $\mathcal{M}$ we denote the ball with center x and radius r by $B_{\mathcal{M}}(x, r)$.

Theorem 11. *$\mathcal{G}_0$ is separable.*

proof. The set of all finite graphs with rational edge weights is a dense subset of $\mathcal{G}_0$. ■

Theorem 12. *$\mathcal{G}_0$ is not locally compact.*

proof. Suppose the contrary, so there is an open neighbourhood $B_{\mathcal{G}_0}(0, r)$ such that $\overline{B_{\mathcal{G}_0}(0, r)}$ is compact. Consider the sequence $\{G_n\}$ in which G_n is the jointless graph with n edges of weight $\frac{r}{n}$. This sequence must have a convergent subsequence. The tiny edges of this subsequence say that it converges to 0. On the other hand the norm of its members are always equal to r implying that the norm of the limit graph must be r, which is a contradiction. ■

Lemma 9. *W_0^s is path connected.*

proof. Take the standard graph w. The function $f : [0, 1] \to W_0^s$, $f(t) = tw$ is a path between 0 and w, so every point is connected to 0 via a path. ■

Theorem 13. *$\mathcal{G}_0$ is path connected.*

proof. The onto function $w \to \widetilde{w}$ from W_0^s to $\mathcal{G}_0$ is continuous and hence takes path connected to path connected. ■

Theorem 14. *$\mathcal{G}_0$ is locally path connected.*

proof. First we show that every ball in W_0^s is path connected. Let $w \in W_0^s$, $r > 0$ and $w' \in B_{W_0^s}(w, r)$. Choose a finite graph w'' from $B_{W_0^s}(w, r)$ (finite graphs are dense in W_0^s). $B_{W_0}(w, r)$ is convex and the graphs w' and w'' have infinitely many isolated vertices in common, so we can define the function $f : [0, 1] \to B_{W_0^s}(w, r)$, $f(t) = tw' + (1 - t)w''$ which is a path between w' and w''. So every member of $B_{W_0^s}(w, r)$ is connected to w'' via a path, therefore the ball is path connected. Now according to the facts that for each $w \in W_0$ we have $B_{\mathcal{G}_0}(\widetilde{w}, r) = \widetilde{B_{W_0^s}(w, r)}$ (which is not difficult to prove) and that the function $w \to \widetilde{w}$ from W_0 to $\mathcal{G}_0$ is continuous, every ball in $\mathcal{G}_0$ is path connected and so $\mathcal{G}_0$ is locally path connected. ■

Acknowledgments. We are grateful to professor Sayyad Ebadollah Mahmoodian, for his valuable advices and comments during this work. We also thank all friends that directly and indirectly gave us hands in the preparation of this paper.

References

[1] Binary codes capable of correcting deletions, insertions, and reversals by: Vladimir I Levenshtein Soviet Physics Doklady, Vol. 10, No. 8. (1966), pp. 707-710.

[2] J. Balogh R. Martin, Edit distance and its computation, Electronic Journal of Combinatorics, 15(1), Research Paper 20, 27 pp