



On a Weighted Generalization of Kendall's Tau Distance

Albert Bruno Piek^{ID} and Evgeniy Petrov^{ID}

Abstract. We introduce a metric on the set of permutations of given order, which is a weighted generalization of Kendall's τ rank distance and study its properties. Using the edge graph of a permutohedron, we give a criterion which guarantees that a permutation lies metrically between another two fixed permutations. In addition, the conditions under which four points from the resulting metric space form a pseudolinear quadruple were found.

Mathematics Subject Classification. 54E35, 05A05, 05C35.

Keywords. Kendall's tau distance, Metric space, Permutation, Permutohedron.

1. Introduction

The concept of metric spaces appeared more than one century ago in works of Maurice Fréchet and Felix Hausdorff. Recall that a metric on a set X is a function $d: X \times X \rightarrow \mathbb{R}_+$, $\mathbb{R}_+ = [0, \infty)$, such that the following conditions hold:

- (i) $d(x, y) = d(y, x)$ (symmetry),
- (ii) $(d(x, y) = 0) \Leftrightarrow (x = y)$ (identity of indiscernibles),
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality),

for all $x, y, z \in X$. The pair (X, d) is called a metric space.

It is possible to define metrics not only on the “abstract” sets of points but also on the sets of various mathematical objects. The well-known Encyclopedia of Distances [7] contains a large number of distances including metrics on such objects as graphs, matrices, strings, permutations, etc. These distances are

Evgeniy Petrov was supported by H2020-MSCA-RISE-2014, Project number 645672 (AM-MODIT: “Approximation Methods for Molecular Modelling and Diagnosis Tools”).

not only interesting to be studied from a theoretical viewpoint, but have also importance in applications.

Many real life settings require comparisons between pairs of objects that cannot be described simply by the Euclidean metric. As an example, the question how to compare different rankings arose in many psychological works, where two or more different observers had the task to rank a set of objects for certain properties. The search for a quantification of the similarity of rankings led to the famous Kendall's rank correlation coefficient introduced in [12]. The Kendall τ distance arises naturally from the rank correlation coefficient and counts the number of pairwise disagreements between two permutations. It is an example of a metric defined on the set S_n of all permutations or, equivalently, ranking lists of order n , see (3). A historical review of Kendall's τ and related coefficients can be found in [13].

The first known to authors weighted generalization of the Kendall τ metric was introduced in [17, 18] as follows:

$$K_w(\pi, \varphi) = \sum_{1 \leq i < j \leq n} w_i w_j \mathbf{1}_{(\pi_j - \pi_i)(\varphi_j - \varphi_i) < 0}, \quad (1)$$

where $w_i w_j > 0$, and $\pi, \varphi \in S_n$.

This generalization was considered in order to “accommodate” the fact that not all predicates are equally important. The metric K_w is equal to the standard Kendall τ distance if $w_i = 1$ for all $i \in \{1, \dots, n\}$. Without any reference to [17], apparently independently, similar generalizations appeared in [14] and [15, 16]. In the latter case authors were motivated from the weighted Kendall's τ correlation coefficient proposed in [21]. In the literature there exist diverse generalizations and modifications of Kendall's τ distance. It is worth to mention a generalization on partial orders [2], generalizations in terms of transpositions [3, 4], the so-called probabilistic Kendall's tau distance [23, p. 125], a weighted Kendall's distance [9], and Kendall's tau sequence distance [5].

Aside from modifications of Kendall's τ , several other noteworthy distance and similarity measures for permutations, respectively rankings, emerged from applications. Another well-known proximity measure for rankings is Spearman's ρ rank coefficient, introduced in [22]. More recent specialized approaches can be found, e.g., in the application fields of medical diagnostics [11], hydrology [10], physiology [24], and neurophysiology [20]. An overview over metrics on S_n is given in [6].

In this paper, using a matrix of weights W , we consider a more general than (1), but for convenience normalized, distance d_W : instead of multiplications of weights $w_i w_j$ we use the weights $w_{i,j}$ directly assigned to the pair of coordinates (i, j) , see (2), and study geometric properties of the space (S_n, d_W) . Such generalization is interesting from the theoretical point of view and is more flexible for possible applications.

In Sect. 2 we prove that (S_n, d_W) in general case is a pseudometric space and is a metric space if and only if all weights are positive, which is stated in Theorem 1. Furthermore, we describe several properties of this space and its metric.

The main result of Sect. 3 is presented in Theorem 2, where we provide a criterion when a point from (S_n, d_W) “lies between” another two fixed points from (S_n, d_W) . It is formulated using an edge-graph of a permutohedron of order n .

In Sect. 4 we investigate the occurrence of special type four-point subsets of the space (S_n, d_W) – so called “pseudolinear quadruples”.

At the end of the paper we formulate a conjecture that suggests a characterization of the metric space (S_n, d_W) by certain geometric properties in a sense that all metric spaces (X, d) , $|X| = n!$, with these properties are isometric to (S_n, d_W) for some weight W . We prove it for the case $n = 3$.

2. Definitions and Basic Properties

In this section we introduce a weighted generalization of Kendall's tau distance and study some of its basic properties.

Denote by S_n the set of all permutations of the numbers $1, \dots, n$. For the permutations $\pi = (\pi_1, \dots, \pi_n)$, $\varphi = (\varphi_1, \dots, \varphi_n) \in S_n$ define the discordance indicator by

$$\text{dsc}_{i,j}(\pi, \varphi) = \mathbf{1}_{(\pi_j - \pi_i)(\varphi_j - \varphi_i) < 0}.$$

Clearly, $\text{dsc}_{i,j}(\pi, \varphi)$ is symmetric with respect to i, j and with respect to π, φ . Define the discordance set of π and φ to be

$$\text{dsc}(\pi, \varphi) = \{(i, j) \mid i < j, \text{dsc}_{i,j}(\pi, \varphi) = 1\}.$$

Remark 1. In the special case where φ is the identity permutation id , $\text{dsc}(\pi, id)$ describes the set of inversions, i.e., all index pairs $i < j$ with $\pi_i > \pi_j$. Typically, the notation $I(\pi)$ is chosen for the inversion set, the cardinality $|I(\pi)|$ of the inversion set is called the inversion number of the permutation π and is a well known property of permutations that measures the sortedness. It is related to the sign of a permutation and was first introduced and used in the Cramer's rule for determinants.

Let $W = (w_{i,j}) \in \mathbb{R}_+^{n \times n}$ be a strictly upper triangular weighting matrix with $w_{i,j} \geq 0$. Define a map $d_W : S_n \times S_n \rightarrow [0, 1]$ as follows:

$$d_W(\pi, \varphi) = \frac{\sum_{i,j=1}^n \text{dsc}_{i,j}(\pi, \varphi) \cdot w_{i,j}}{\sum_{i,j=1}^n w_{i,j}}. \quad (2)$$

Everywhere below we consider that $\sum_{i,j=1}^n w_{i,j} \neq 0$. The distance (2) generalizes the normalized Kendall τ ranking distance, which is defined as

$$K(\pi, \varphi) = \frac{|\text{dsc}(\pi, \varphi)|}{n(n-1)/2}. \quad (3)$$

This distance coincidences with $d_W(\pi, \varphi)$ if $W = W^\tau = (w_{i,j})$, where $w_{i,j} = 1$, $i < j$. Thereby, d_{W^τ} is also related to Kendall's τ correlation coefficient $\tau(\pi, \varphi)$ [12] by $\tau(\pi, \varphi) = 1 - 2d_{W^\tau}$.

Recall that a pseudometric space is a generalization of a metric space in which the distance between two distinct points can be zero, i.e., instead of

axiom (ii) in the definition of metric spaces we have the condition $d(x, x) = 0$. In this case d is called a pseudometric.

Theorem 1. *The pair (S_n, d_W) is a pseudometric space, $n \geq 2$. Moreover, d_W is a metric if and only if all the weights $w_{i,j}$ are positive for $i < j$.*

Proof. Symmetry can be seen immediately. Zero distance $d_W(\pi, \pi) = 0$ for equal permutations follows from the equality $\text{dsc}_{i,j}(\pi, \pi) = \mathbf{1}_{(\pi_j - \pi_i)^2 < 0} = 0$. Let $\pi, \varphi, \psi \in S_n$. To prove the triangle inequality consider the difference

$$\Delta = d_W(\pi, \varphi) + d_W(\varphi, \psi) - d_W(\pi, \psi).$$

In order to fulfill the triangle inequality, this difference has to be nonnegative. One can see that Δ can be written as

$$\Delta = \frac{\sum_{i,j=1}^n (\text{dsc}_{i,j}(\pi, \varphi) + \text{dsc}_{i,j}(\varphi, \psi) - \text{dsc}_{i,j}(\pi, \psi)) w_{i,j}}{\sum_{i,j=1}^n w_{i,j}} \geq 0. \quad (4)$$

The summands $\text{dsc}_{i,j}(\pi, \varphi) + \text{dsc}_{i,j}(\varphi, \psi) - \text{dsc}_{i,j}(\pi, \psi)$ in the numerator only takes values in $\{0, 2\}$ since

$$\begin{aligned} (\text{dsc}_{i,j}(\pi, \varphi) = \text{dsc}_{i,j}(\varphi, \psi)) &\Rightarrow (\text{dsc}_{i,j}(\pi, \psi) = 0), \\ (\text{dsc}_{i,j}(\pi, \varphi) \neq \text{dsc}_{i,j}(\varphi, \psi)) &\Rightarrow (\text{dsc}_{i,j}(\pi, \psi) = 1). \end{aligned}$$

Hence and because of the fact that all weights $w_{i,j}$ are nonnegative, inequality (4) evidently holds. Thereby the triangle inequality holds for d_W making it a pseudometric.

Consider now only positive weights $w_{i,j}$ for $i < j$. Then

$$d_W(\pi, \varphi) = 0 \Leftrightarrow \text{dsc}_{i,j}(\pi, \varphi) = 0 \text{ for all } 1 \leq i < j \leq n.$$

Since both permutations are thereby concordant for every index pair i, j , they must be equal and d_W satisfies the identity of indiscernibles. \square

Remark 2. The requirement of positive weights in (2) is sharp in the sense that if $w_{i,j} = 0$ for at least one pair (i, j) one can always find pairs of permutations with distance zero. These can be obtained by swapping the i th and j th element of an arbitrary permutation.

Remark 3. In the case $n = 1$ we have $S_1 = \{(1)\}$ and (S_1, d_W) is a trivial one-point metric space.

Let us define the following permutation for given $\pi = (\pi_1, \dots, \pi_n)$:

$$\hat{\pi} = (n + 1 - \pi_1, \dots, n + 1 - \pi_n) \text{ (ordinal inverse).}$$

The next proposition describes some basic geometric properties of the space (S_n, d_W) .

Proposition 1. *The following conditions hold for every pseudometric space (S_n, d_W) , $n \geq 2$, and for all $\pi, \varphi \in S_n$:*

- (i) The pseudometric d_W is scaling invariant with respect to W , i.e., the equality

$$d_{aW}(\pi, \varphi) = d_W(\pi, \varphi)$$

holds for every $a > 0$, and for every $W \in \mathbb{R}_+^{n \times n}$.

- (ii) The pseudometric d_W is subadditive with respect to W , i.e., the inequality

$$d_{W+V}(\pi, \varphi) \leq d_W(\pi, \varphi) + d_V(\pi, \varphi)$$

holds for all $W, V \in \mathbb{R}_+^{n \times n}$.

- (iii) The equality $d_W(\pi, \hat{\varphi}) = 1 - d_W(\pi, \varphi)$ holds.
 (iv) The equality $d_W(\pi, \varphi) = d_W(\hat{\pi}, \hat{\varphi})$ holds.
 (v) The equality $d_W(\pi, \varphi) \leq 1$ holds. Moreover, if d_W is a metric, then $d_W(\pi, \varphi) = 1$ if and only if $\varphi = \hat{\pi}$.

Proof. The proofs of the statements (i) and (ii) are straightforward and left to the reader. (iii) For each index pair $i < j$ it follows from the definition of the ordinal inverse that

$$\text{sgn}(\hat{\varphi}_j - \hat{\varphi}_i) = \text{sgn}(n + 1 - \varphi_j - n - 1 + \varphi_i) = \text{sgn}(\varphi_i - \varphi_j) = -\text{sgn}(\varphi_j - \varphi_i).$$

In consequence, $\text{dsc}_{i,j}(\pi, \hat{\varphi}) = \mathbf{1}_{(\pi_j - \pi_i)(\varphi_j - \varphi_i) > 0} = 1 - \text{dsc}_{i,j}(\pi, \varphi)$. Since this statement holds for every index pair $i < j$, the result remains when the weighted and normalized sum over all pairs is considered, leading to (iii). (iv) Using statement (iii) we have $d_W(\hat{\pi}, \hat{\varphi}) = 1 - d_W(\hat{\pi}, \varphi) = 1 - d_W(\varphi, \hat{\pi}) = 1 - (1 - d_W(\varphi, \pi)) = d_W(\pi, \varphi)$. (v) The inequality $d_W(\pi, \varphi) \leq 1$ follows directly from (2). Let d_W be a metric and let $\varphi = \hat{\pi}$. Using conditions (iv) and (iii) for the proof consider the sequence of equivalences: $\varphi = \hat{\pi}$ iff $d_W(\hat{\pi}, \varphi) = 0$ iff $d_W(\pi, \hat{\varphi}) = 0$ iff $1 - d_W(\pi, \varphi) = 0$ iff $d_W(\pi, \varphi) = 1$. \square

3. Betweenness of Points in (S_n, d_W)

In this section we continue to study geometric properties of the space (S_n, d_W) by characterizing triplets of points from S_n which satisfy the ternary relation “to lie between”. This relation is intuitive for points belonging to some straight line, plane or three-dimensional space. Menger [19, p. 77] seems to be the first who formulated the concept of “metric betweenness” for general metric spaces. Let (X, d) be a metric space, and let x, y and z be different points from X . The point y lies between x and z , if $d(x, z) = d(x, y) + d(y, z)$. This concept is used also in the present time for the study of metric spaces, see, e.g., [1].

Recall that an undirected graph is a pair (V, E) consisting of a nonempty set V and a (probably empty) set E whose elements are unordered pairs of different points from V . For a graph $G = (V, E)$, the sets $V = V(G)$ and $E = E(G)$ are called the set of vertices and the set of edges, respectively. A path in a graph G is a subgraph P of G for which

$$V(P) = \{x_0, \dots, x_k\}, \quad E(P) = \{\{x_0, x_1\}, \dots, \{x_{k-1}, x_k\}\},$$

where all x_i are distinct. Sometimes for convenience we refer to a path by the natural sequence of its vertices, say, $P = \{x_0, \dots, x_k\}$. A finite graph C is a

cycle if $|V(C)| \geq 3$ and there exists an enumeration (v_1, \dots, v_n) of its vertices such that

$$(\{v_i, v_j\} \in E(C)) \Leftrightarrow (|i - j| = 1 \text{ or } |i - j| = n - 1).$$

A cycle is simple if no repetitions of vertices and edges allowed.

Denote by $\text{tr}_{ij}(\pi)$ a permutation which is obtained from $\pi = \{\pi_1, \dots, \pi_n\}$ by transposition of two elements π_i and π_j , $i \neq j$, $i, j \in \{1, \dots, n\}$, only in the case if

$$|\pi_i - \pi_j| = 1. \quad (5)$$

Let $G_n = G_n(V, E)$ be an undirected graph such that $V(G_n) = S_n$ and $\{\pi, \varphi\} \in E(G_n)$ if and only if $\varphi = \text{tr}_{ij}(\pi)$ for some $i \neq j$. The graph G_n is known as an edge-graph of a permutohedron of order n . Let us remember that an adjacent transposition is a transposition $(\pi_i \pi_j)$ where the two elements are consecutive, i.e., when equality (5) holds. In other words, two vertices of the graph G_n are connected by the edge if and only if one vertex is obtained from the other by applying an adjacent transposition. For the edges $E(G_n)$ we define a labeling function $l: E(G_n) \rightarrow \{(i, j) \mid i < j\}$ in the following way: let e be the edge $\{\pi, \varphi\}$, then it is labeled with $l(e) = (i, j)$, which is the index pair for which $\varphi = \text{tr}_{ij}(\pi)$ holds. The labeled graph G_4 is depicted at Fig. 1.

Recall that a permutohedron of order n is an $(n-1)$ -dimensional polytope embedded in an n -dimensional Euclidean space, which is a convex hull of all $n!$ points formed by permuting the coordinates of the vector $(1, 2, \dots, n)$. Furthermore, the permutohedron is vertex-transitive, i.e., for every $\tau \in S_n$ the following implication holds:

$$\{\pi, \varphi\} \in E(G_n) \Rightarrow \{\pi \circ \tau, \varphi \circ \tau\} \in E(G_n).$$

This is indeed true, since $\varphi = \text{tr}_{ij}(\pi)$ implies $\varphi \circ \tau = \text{tr}_{\tau_i, \tau_j}(\pi \circ \tau)$, as τ only reorders the indices. In consequence, permutohedra are highly symmetric, which can be seen in the visualizations of G_n for the cases $n = 4$ and $n = 3$. Both are depicted in Figs. 1 and 2, respectively.

Recall that a distance d_G between two vertices in a connected graph G is the number of edges in a shortest path connecting them. The distance d_G satisfies axioms (i)–(iii) of a metric space. Thus, d_G is metric induced by G . In general, the shortest path connecting two any different vertices of G_n is not obligatory unique. For the cases $n = 1$ and $n = 2$, the graphs are given by $V(G_1) = \{(1)\}$, $E(G_1) = \emptyset$ and $V(G_2) = \{(1, 2), (2, 1)\}$, $E(G_2) = \{\{(1, 2), (2, 1)\}\}$ respectively.

Proposition 2. *For any $n \geq 1$ and for any $\pi, \varphi \in S_n$ the distance d_{G_n} between these permutations equals the number of discordant pairs $|\text{dsc}(\pi, \varphi)|$.*

Proof. For $n = 1, 2$ the statement is trivial. Let $n \geq 3$. An adjacent transposition on a discordant pair decreases the inversion number by exactly one. Indeed, in this case only two consecutive integers are swapped and all other elements of permutation preserve their order in relation to each of the swapped elements. So every path between π and φ require at least $|\text{dsc}(\pi, \varphi)|$ edges.

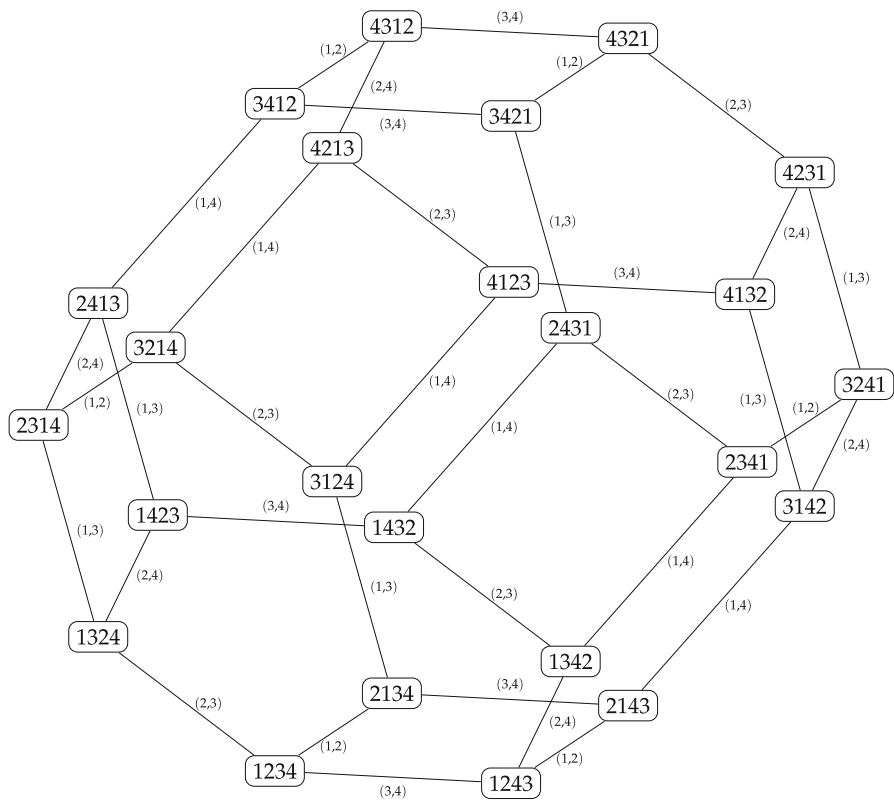


FIGURE 1. The labeled undirected graph G_4

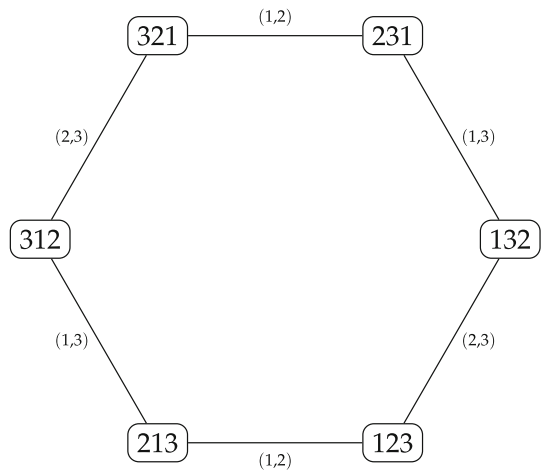


FIGURE 2. The labeled undirected graph G_3

Let us show the existence of such path by induction on $|\text{dsc}(\pi, \varphi)|$. Since the graph is vertex transitive we may assume that $\pi = id$. If $|\text{dsc}(id, \varphi)| = 0$, then $\varphi = id$ and $d_{G_n}(id, id) = 0$. Let $\varphi = (\varphi_1, \dots, \varphi_n)$ and let $|\text{dsc}(id, \varphi)| \neq 0$. Hence there exists at least one pair of elements φ_i and φ_j such that $i < j$ and $\varphi_i - \varphi_j = 1$. Indeed, assuming the opposite we immediately get that 1 must be before 2, 2 before 3, ... and, in consequence, $\varphi = id$. Let $\hat{\varphi} = \text{tr}_{ij}(\varphi)$. Then $|\text{dsc}(id, \hat{\varphi})| = |\text{dsc}(id, \varphi)| - 1$. By induction hypothesis $d_{G_n}(id, \hat{\varphi}) = |\text{dsc}(id, \hat{\varphi})|$. Since $\text{dsc}(id, \varphi) = \text{dsc}(id, \hat{\varphi}) \cup \{(i, j)\}$ we obtain the necessary equality. \square

Let $\mathcal{P} = \{\pi = \psi_1, \dots, \psi_n = \varphi\}$ be a path joining π and φ in G_n . By the definition of G_n we have

$$\psi_2 = \text{tr}_{i_1, j_1}(\psi_1) \quad \text{and} \quad \text{dsc}(\psi_1, \psi_2) = \{(i_1, j_1)\},$$

$$\dots$$

$$(6)$$

$$\psi_n = \text{tr}_{i_{n-1}, j_{n-1}}(\psi_{n-1}) \quad \text{and} \quad \text{dsc}(\psi_{n-1}, \psi_n) = \{(i_{n-1}, j_{n-1})\}.$$

The following corollary follows directly from Proposition 2 and the definition of the graph G_n .

Corollary 1. *Let π, φ be different vertices of G_n and let \mathcal{P} be a path in G_n connecting π and φ . Then \mathcal{P} is a shortest-path if and only if all the pairs $(i_1, j_1), \dots, (i_{n-1}, j_{n-1})$ defined by (6) are different.*

Proposition 3. *Let π, φ be different vertices of the graph G_n , $n \geq 2$. Then the following statements are equivalent for every $\psi \in V(G_n)$:*

- (i) $\text{dsc}(\pi, \varphi) = \text{dsc}(\pi, \psi) \cup \text{dsc}(\psi, \varphi)$.
- (ii) $\text{dsc}(\pi, \psi) \cap \text{dsc}(\psi, \varphi) = \emptyset$.
- (iii) *There exists a shortest-path $\mathcal{P}_{\pi, \varphi}$ between π and φ such that $\psi \in V(\mathcal{P}_{\pi, \varphi})$.*

Proof. Let us prove the implication (i) \Rightarrow (ii) by contradiction. Let us assume that there exists a pair

$$(i, j) \in \text{dsc}(\pi, \psi) \cap \text{dsc}(\psi, \varphi). \quad (7)$$

Since $\text{dsc}(\pi, \psi) \cap \text{dsc}(\psi, \varphi) \subseteq \text{dsc}(\pi, \psi) \cup \text{dsc}(\psi, \varphi)$, by (i) we have $(i, j) \in \text{dsc}(\pi, \varphi)$. From (7) it follows that $\text{dsc}_{i,j}(\pi, \psi) = \text{dsc}_{i,j}(\psi, \varphi) = 1$ or equally

$$(\pi_j - \pi_i)(\psi_j - \psi_i) < 0 \quad \text{and} \quad (\psi_j - \psi_i)(\varphi_j - \varphi_i) < 0.$$

Multiplying both left sides gives

$$\begin{aligned} (\pi_j - \pi_i)(\psi_j - \psi_i)^2(\varphi_j - \varphi_i) &> 0 \\ \Rightarrow (\pi_j - \pi_i)(\varphi_j - \varphi_i) &> 0 \\ \Rightarrow \text{dsc}_{i,j}(\pi, \varphi) &= 0 \\ \Rightarrow (i, j) &\notin \text{dsc}(\pi, \varphi). \end{aligned}$$

This contradicts to our assumption.

(ii) \Rightarrow (iii) Let $\mathcal{P}_{\pi, \psi}$ and $\mathcal{P}_{\psi, \varphi}$ be any shortest paths between the respective vertices and let $\mathcal{P}_{\pi, \varphi} = \mathcal{P}_{\pi, \psi} \cup \mathcal{P}_{\psi, \varphi}$ be the compound path from π to φ over ψ . By (ii) all the pairs $(i_1, j_1), \dots, (i_{n-1}, j_{n-1})$ defined by (6) for the path $\mathcal{P}_{\pi, \varphi}$ are different. Therefore by Corollary 1 the compound $\mathcal{P}_{\pi, \varphi}$ is a shortest path between π and φ .

(iii) \Rightarrow (i) Let the relations (6) hold for $\mathcal{P}_{\pi, \varphi}$. By Corollary 1 all the pairs $(i_1, j_1), \dots, (i_{n-1}, j_{n-1})$ are different. From (5) it follows that

$$\text{dsc}(\psi_1, \psi_{k+1}) = \text{dsc}(\psi_1, \psi_k) \cup \{(i_k, j_k)\}, \quad k = 1, \dots, n-1.$$

Hence,

$$\begin{aligned} \text{dsc}(\pi, \psi_k) &= \text{dsc}(\psi_1, \psi_k) = \{(i_1, j_1), \dots, (i_{k-1}, j_{k-1})\}, \quad k = 2, \dots, n-1, \\ \text{dsc}(\pi, \varphi) &= \text{dsc}(\psi_1, \psi_n) = \{(i_1, j_1), \dots, (i_{n-1}, j_{n-1})\}. \end{aligned}$$

Analogously,

$$\text{dsc}(\psi_k, \varphi) = \{(i_k, j_k), \dots, (i_{n-1}, j_{n-1})\} \quad k = 2, \dots, n-1,$$

which establishes (i) with $\psi = \psi_k$. \square

Theorem 2. For any $n \geq 3$ and any weighting matrix W (strictly upper triangular and positive), and any three different permutations $\pi, \psi, \varphi \in S_n$, ψ lies between π and φ with respect to the metric d_{G_n} if and only if ψ lies between π and φ with respect to d_W .

Proof. Let ψ lie between π and φ with respect to d_{G_n} . Then, ψ belongs to some shortest path $\mathcal{P}_{\pi, \varphi}$ connecting π and φ in G_n . The distance d_W is the normalized sum of weights associated to discordant pairs between two permutations. These discordant pairs are exactly the labels of edges on shortest paths between those two permutations. By Corollary 1 the labels on the shortest-path $\mathcal{P}_{\pi, \varphi}$ are exactly the disjoint union of the labels on the shortest-paths $\mathcal{P}_{\pi, \psi}$ and $\mathcal{P}_{\psi, \varphi}$. Hence we have the equality $d_W(\pi, \varphi) = d_W(\pi, \psi) + d_W(\psi, \varphi)$.

Let us show the converse implication by contradiction. Let ψ be a permutation that does not lie between π and φ with respect to d_{G_n} . Then ψ does not belong to any shortest path connecting π and φ in G_n . Proposition 3 implies $\text{dsc}(\pi, \varphi) \neq \text{dsc}(\pi, \psi) \cup \text{dsc}(\psi, \varphi)$. Now if $\text{dsc}_{i,j}(\pi, \varphi) = 1$, there are two possibilities:

$$\text{dsc}_{i,j}(\pi, \psi) = 1 \quad \text{and} \quad \text{dsc}_{i,j}(\psi, \varphi) = 0$$

or

$$\text{dsc}_{i,j}(\pi, \psi) = 0 \quad \text{and} \quad \text{dsc}_{i,j}(\psi, \varphi) = 1.$$

In other words $(i, j) \in \text{dsc}(\pi, \varphi)$ implies $(i, j) \in \text{dsc}(\pi, \psi) \cup \text{dsc}(\psi, \varphi)$. Hence, $\text{dsc}(\pi, \varphi)$ is a proper subset of $\text{dsc}(\pi, \psi) \cup \text{dsc}(\psi, \varphi)$ and the equality

$$\frac{\sum_{i,j=1}^n \text{dsc}_{i,j}(\pi, \varphi) \cdot w_{i,j}}{\sum_{i,j=1}^n w_{i,j}} = \frac{\sum_{i,j=1}^n \text{dsc}_{i,j}(\pi, \psi) \cdot w_{i,j}}{\sum_{i,j=1}^n w_{i,j}} + \frac{\sum_{i,j=1}^n \text{dsc}_{i,j}(\psi, \varphi) \cdot w_{i,j}}{\sum_{i,j=1}^n w_{i,j}}$$

is impossible, since d_W is a metric and all $w_{i,j}$ are positive. \square

4. Pseudolinear Quadruples in (S_n, d_W)

The aim of this section is to describe the occurrence of special type four-point subsets of the space (S_n, d_W) , the so called “pseudolinear quadruples”.

In 1928 Menger [19] proved that if every three points of a metric space X , $|X| \geq 3$, are embeddable into \mathbb{R}^1 , then X is isometric to some subset of \mathbb{R}^1

or X is a pseudolinear quadruple. Recall that a four-point metric space (X, d) is called a pseudolinear quadruple if there exists an enumeration x_1, x_2, x_3, x_4 of the points of X such that the equalities

$$\begin{aligned} d(x_1, x_2) &= d(x_3, x_4) = s, \quad d(x_2, x_3) = d(x_4, x_1) = t, \\ d(x_2, x_4) &= d(x_3, x_1) = s + t \end{aligned} \quad (8)$$

hold with some positive reals s and t . Note also that equilateral pseudolinear quadruples are known by their extremal properties [8].

Let (X, d) be a metric space. Recall that for every nonempty set $A \subseteq X$ the quantity

$$\text{diam}A = \sup\{d(x, y) : x, y \in A\}$$

is the diameter of A . We shall say that points a, b are diametrical for the set A if $d(a, b) = \text{diam}A$.

Everywhere below in this section we consider that $n \geq 3$, since this is a necessary condition for the existence of pseudolinear quadruples in (S_n, d_W) .

Proposition 4. *Let π and φ be nondiametrical points in the pseudometric space (S_n, d_W) . Then the set $X = \{\pi, \varphi, \hat{\pi}, \hat{\varphi}\}$ forms a pseudolinear quadruple.*

Proof. By condition (iv) of Proposition 1 we have

$$d_W(\pi, \varphi) = d_W(\hat{\pi}, \hat{\varphi}), \quad d_W(\pi, \hat{\varphi}) = d_W(\varphi, \hat{\pi}).$$

Using conditions (iii) and (v) of the same proposition we obtain the following equalities:

$$\begin{aligned} d_W(\pi, \varphi) + d_W(\varphi, \hat{\pi}) &= d_W(\pi, \varphi) + 1 - d_W(\varphi, \pi) = 1 = d_W(\pi, \hat{\pi}), \\ d_W(\varphi, \pi) + d_W(\pi, \hat{\varphi}) &= d_W(\varphi, \pi) + 1 - d_W(\pi, \varphi) = 1 = d_W(\varphi, \hat{\varphi}). \end{aligned}$$

Thus, (X, d_W) is a pseudolinear quadruple with $x_1 = \pi$, $x_2 = \varphi$, $x_3 = \hat{\pi}$, $x_4 = \hat{\varphi}$. \square

It is easy to see that every cycle in G_n is even. Let C be a labeled cycle in G_n . We shall say that C has a symmetric labeling if $l(e) = l(\bar{e})$, where \bar{e} is an edge opposite to e in C . Denote by $E_C(i, j)$ the set of edges of a cycle C labeled by the label (i, j) .

Proposition 5. *Let C be a simple cycle in G_n having a symmetric labeling. Let for every label (i, j) of the cycle C the equality $|E_C(i, j)| = 2(2k - 1)$ hold for some $k \in \mathbb{N}^+$. Then for every different non opposite vertices π, φ of C the set $\{\pi, \varphi, \bar{\pi}, \bar{\varphi}\}$ form a pseudolinear quadruple in (S_n, d_W) , where $\bar{\pi}, \bar{\varphi}$ are opposite vertices to π, φ in C , respectively.*

Proof. Let $\pi \in V(C)$ and let $\mathcal{P}_{\pi, \bar{\pi}}$ be one of the paths connecting π and $\bar{\pi}$ in C . Without loss of generality, consider that $\varphi \in V(\mathcal{P}_{\pi, \bar{\pi}})$. Since C has a symmetric labeling, for every label (i, j) of C the number of edges labeled by (i, j) and belonging to $\mathcal{P}_{\pi, \bar{\pi}}$ is odd. Hence, the number of edges labeled by (i, j) and belonging to $\mathcal{P}_{\pi, \varphi}$ ($\mathcal{P}_{\varphi, \bar{\pi}}$) is odd (even) or vice versa. Thus, $\text{dsc}_{i,j}(\pi, \varphi) = 1$ ($\text{dsc}_{i,j}(\varphi, \bar{\pi}) = 0$) or vice versa and $\text{dsc}_{i,j}(\pi, \bar{\pi}) = 1$. Anyway,

$$\text{dsc}_{i,j}(\pi, \varphi) + \text{dsc}_{i,j}(\varphi, \bar{\pi}) = \text{dsc}_{i,j}(\pi, \bar{\pi}) \quad (9)$$

for every label (i, j) of the graph C . Analogously,

$$\text{dsc}_{i,j}(\pi, \varphi) + \text{dsc}_{i,j}(\pi, \bar{\varphi}) = \text{dsc}_{i,j}(\varphi, \bar{\varphi}). \quad (10)$$

Equalities (9), (10) and (2) give

$$\begin{aligned} d_W(\pi, \varphi) + d_W(\varphi, \bar{\pi}) &= d_W(\pi, \bar{\pi}), \\ d_W(\pi, \varphi) + d_W(\pi, \bar{\varphi}) &= d_W(\varphi, \bar{\varphi}). \end{aligned}$$

By symmetric labeling of C we have

$$\text{dsc}(\pi, \varphi) = \text{dsc}(\bar{\pi}, \bar{\varphi}), \quad \text{dsc}(\varphi, \bar{\pi}) = \text{dsc}(\pi, \bar{\varphi}).$$

Hence, by (2)

$$d_W(\pi, \varphi) = d_W(\bar{\pi}, \bar{\varphi}), \quad d_W(\varphi, \bar{\pi}) = d_W(\pi, \bar{\varphi}).$$

Thus, equalities (8) are satisfied with

$$\begin{aligned} x_1 &= \pi, \quad x_2 = \varphi, \quad x_3 = \bar{\pi}, \quad x_4 = \bar{\varphi}, \quad \text{and} \\ s &= d_W(\pi, \varphi) = d_W(\bar{\pi}, \bar{\varphi}), \quad t = d_W(\varphi, \bar{\pi}) = d_W(\pi, \bar{\varphi}). \end{aligned}$$

□

In the case $k = 1$ Proposition 5 implies the following.

Corollary 2. *Let C be a simple cycle in G_n having a symmetric labeling and let for every $\pi \in V(C)$ different edges of the path $\mathcal{P}_{\pi, \bar{\pi}} \subseteq C$ have different labels, where $\bar{\pi}$ is a vertex opposite to π in C . Then for every different nonopposite vertices π, φ of C the set $\{\pi, \varphi, \bar{\pi}, \bar{\varphi}\}$ form a pseudolinear quadruple in (S_n, d_W) , where $\bar{\pi}, \bar{\varphi}$ are opposite vertices to π, φ in C , respectively.*

Remark 4. The assertion converse to Corollary 2 does not hold. Consider the permutations

$$\alpha = (1, 2, 3, 4), \quad \beta = (4, 1, 2, 3), \quad \gamma = (4, 2, 3, 1), \quad \delta = (1, 3, 4, 2).$$

We have

$$\begin{aligned} \text{dsc}(\alpha, \beta) &= \{(1, 2), (1, 3), (1, 4)\} = \text{dsc}(\gamma, \delta), \\ \text{dsc}(\beta, \gamma) &= \{(2, 4), (3, 4)\} = \text{dsc}(\alpha, \delta), \\ \text{dsc}(\alpha, \gamma) &= \{(1, 2), (1, 3), (1, 4), (2, 4), (3, 4)\} = \text{dsc}(\beta, \delta). \end{aligned}$$

This implies that $(\alpha, \beta, \gamma, \delta)$ is a pseudolinear quadruple in (S_4, d_W) . Let us show that this pseudolinear quadruple is not a part of a symmetric labeled cycle in G_4 . Therefore, we show that there are no paths $P_{\alpha, \beta}$ from α to β and $P_{\gamma, \delta}$ from γ to δ in G_4 such that they have the same length and the same labeling. Denote by $l_a(\pi)$ the set of all labels of the edges adjacent to π . One can see from Fig. 1 that

$$\begin{aligned} l_a(\alpha) &= \{(1, 2), (2, 3), (3, 4)\}, \\ l_a(\gamma) &= \{(2, 4), (2, 3), (1, 3)\}. \end{aligned}$$

Hence, $l_a(\alpha) \cap l_a(\gamma) = \{(2, 3)\}$. For the next point α_1 on $P_{\alpha, \beta}$ and the next point γ_1 on $P_{\gamma, \delta}$ the labels must be equal, therefore $l(\{\alpha, \alpha_1\}) = l(\{\gamma, \gamma_1\}) = (\{2, 3\})$. Consequently, $\alpha_1 = (1, 3, 2, 4)$, $\gamma_1 = (4, 3, 2, 1)$. Again

$$\begin{aligned} l_a(\alpha_1) &= \{(2, 3)(1, 4), (2, 4)\}, \\ l_a(\gamma_1) &= \{(2, 3)(1, 2), (3, 4)\}, \end{aligned}$$

and $l_a(\alpha_1) \cap l_a(\gamma_1) = \{(2, 3)\}$. Thus, there is no other way than backwards for the labels to be symmetric. In conclusion, there are no symmetrically labeled paths $P_{\alpha, \beta}, P_{\gamma, \delta}$ and the pseudolinear quadruple $(\alpha, \beta, \gamma, \delta)$ lies on no symmetric labeled cycle in G_4 .

Example 1. Let us show an example of a cycle satisfying condition of Proposition 5 with $k > 1$. Let $\pi_1 = (1, 2, 3, 4, 5, 6, 7, 8) \in S_8$ and let $C = (\pi_1, \dots, \pi_{12})$ such that

$$\begin{aligned} &\pi_1(1, 2)\pi_2(3, 4)\pi_3(1, 2)\pi_4(5, 6)\pi_5(1, 2)\pi_6(7, 8)\pi_7 \\ &\pi_7(1, 2)\pi_8(3, 4)\pi_9(1, 2)\pi_{10}(5, 6)\pi_{11}(1, 2)\pi_{12}(7, 8)\pi_1. \end{aligned}$$

Here and below $\pi_k(i, j)\pi_l$ means that the permutation π_l is obtained from π_k by transposition i -th and j -th elements.

Example 2. Let us show that a symmetric labeling of a cycle C in G_n is not sufficient for every four points $\{\pi, \varphi, \bar{\pi}, \bar{\varphi}\}$ of this cycle forming a pseudolinear quadruple in (S_n, d_W) , where $\bar{\pi}, \bar{\varphi}$ are opposite vertices to π, φ in C , respectively. Indeed, let $\pi_1 = (1, 2, 3, 4, 5, 6) \in S_6$ and let $C = (\pi_1, \dots, \pi_8)$ such that

$$\begin{aligned} &\pi_1(1, 2)\pi_2(3, 4)\pi_3(1, 2)\pi_4(5, 6)\pi_5 \\ &\pi_5(1, 2)\pi_6(3, 4)\pi_7(1, 2)\pi_8(5, 6)\pi_1. \end{aligned}$$

Consider a quadruple of points $\{\pi_1, \pi_3, \pi_5, \pi_7\}$. For these points holds that $\text{dsc}(\pi_1, \pi_3) = \{(1, 2), (3, 4)\}$, $\text{dsc}(\pi_3, \pi_5) = \{(1, 2), (5, 6)\}$ and $\text{dsc}(\pi_1, \pi_5) = \{(3, 4), (5, 6)\}$. Using (2) we see that neither of the pairs $\{\pi_1, \pi_3\}$, $\{\pi_3, \pi_5\}$, $\{\pi_1, \pi_5\}$ can be a diametrical pair of the pseudolinear quadruple $\{\pi_1, \pi_3, \pi_5, \pi_7\}$.

By $L(\mathcal{P})$ we denote below the set of all labels of the path \mathcal{P} and by $\mathcal{P}_{\pi, \varphi}$ any of the shortest paths between π and φ in G_n . Let $I(W)$ be the set of all elements of the matrix W which lie above the main diagonal, i.e., $I(W) = \{w_{i, j}\}_{i < j}$.

Proposition 6. *Let (S_n, d_W) be a metric space, i.e., $w_{i, j} > 0$ for all $i < j$, W be a weight such that for every two different subsets $S_1, S_2 \subseteq I(W)$ the relation*

$$\sum_{r_i \in S_1} r_i \neq \sum_{r_i \in S_2} r_i, \quad (11)$$

holds and let $\pi, \varphi, \bar{\pi}, \bar{\varphi}$ be pairwise different points in S_n . Then the following conditions are equivalent:

- (i) *The set $\{\pi, \varphi, \bar{\pi}, \bar{\varphi}\}$ form a pseudolinear quadruple in (S_n, d_W) with the diameter $d_W(\pi, \bar{\pi}) = d_W(\varphi, \bar{\varphi})$.*

$$(ii) \quad L(\mathcal{P}_{\pi, \varphi}) = L(\mathcal{P}_{\bar{\pi}, \bar{\varphi}}), \quad L(\mathcal{P}_{\varphi, \bar{\pi}}) = L(\mathcal{P}_{\bar{\varphi}, \pi}), \quad L(\mathcal{P}_{\pi, \varphi}) \cap L(\mathcal{P}_{\varphi, \bar{\pi}}) = \emptyset.$$

Proof. The implication (ii) \Rightarrow (i) is almost evident for any W .

Let us prove the implication (i) \Rightarrow (ii) by contradiction. Without loss of generality suppose first that $L(\mathcal{P}_{\pi, \varphi}) \neq L(\mathcal{P}_{\bar{\pi}, \bar{\varphi}})$. Hence, $\text{dsc}(\pi, \varphi) \neq \text{dsc}(\bar{\pi}, \bar{\varphi})$. Using (2) and (11) it follows that the equality $d_W(\pi, \varphi) = d_W(\bar{\pi}, \bar{\varphi})$ is impossible.

Suppose that $L(\mathcal{P}_{\pi, \varphi}) \cap L(\mathcal{P}_{\varphi, \bar{\pi}}) \neq \emptyset$ then there exists a label l such that $l \in L(\mathcal{P}_{\pi, \varphi}) \cap L(\mathcal{P}_{\varphi, \bar{\pi}})$. Since $\mathcal{P}_{\pi, \varphi} \cup \mathcal{P}_{\varphi, \bar{\pi}}$ is a path connecting π and $\bar{\pi}$ in G_n and the label l appears twice in this path, we have that $l \notin \text{dsc}(\pi, \bar{\pi})$. Clearly, $l \in \text{dsc}(\pi, \varphi)$, $l \in \text{dsc}(\varphi, \bar{\pi})$. Again, using only (2) we see that the equality $d_W(\pi, \varphi) + d_W(\varphi, \bar{\pi}) = d_W(\pi, \bar{\pi})$ is impossible. \square

Remark 5. There is a simple combinatorial description of all faces of a permutohedron of order n : its k -faces correspond to ordered partitions of the set $\{1, \dots, n\}$ into $n - k$ nonempty parts [25]. Proposition 4 describes pseudolinear quadruples with the diameter 1. It is possible to show that every subset of G_n formed from the vertices of k -faces, $k \geq 2$, contains a cycle C , satisfying conditions of Corollary 2. In other words every k -face $2 \leq k \leq n - 2$, contains pseudolinear quadruples in X with diameter strictly less than $\text{diam}X = 1$.

Conjecture 1. Let (X, d) be a finite metric space such that the following conditions hold:

- (i) $|X| = n!$;
- (ii) For every $x \in X$ there is a unique $\bar{x} \in X$ such that $d(x, \bar{x}) = \text{diam}X$;
- (iii) For every two non-diametrical points x, y the set $\{x, y, \bar{x}, \bar{y}\}$ form a pseudolinear quadruple;
- (iv) For every two different points $x, y \in X$ there exists $z \in X$ and a sequence of points $z = p_0, p_1, \dots, p_k = \bar{z}$ such that $x, y \in \{p_0, p_1, \dots, p_k\}$ and for every $0 \leq i < j \leq k$ the equality

$$d(p_i, p_j) = d(p_i, p_{i+1}) + \dots + d(p_{j-1}, p_j)$$

holds, where $k = \binom{n}{2}$. For $k > \binom{n}{2}$ such sequences do not exist.

Clearly, for every metric space (S_n, d_W) conditions (i)–(iv) hold. Thus this conjecture asserts that these conditions completely define the structure of (S_n, d_W) up to the weight W .

Proof. (For the case $n = 3$.) Let (X, d) be a metric space satisfying conditions (i)–(iv). And let $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. By condition (ii), without loss of generality, consider that

$$d(x_1, x_4) = d(x_2, x_5) = d(x_3, x_6) = \text{diam}X.$$

It follows from (iii) that

$$\begin{aligned} d(x_2, x_4) &= d(x_1, x_5) = d, & d(x_2, x_6) &= d(x_3, x_5) = e, \\ d(x_1, x_3) &= d(x_4, x_6) = f, & d(x_1, x_2) &= d(x_4, x_5) = a, \\ d(x_2, x_3) &= d(x_5, x_6) = b, & d(x_3, x_4) &= d(x_6, x_1) = c, \end{aligned}$$

and

$$\text{diam}X = a + d = b + e = c + f, \quad (12)$$

see Fig. 3. Clearly, from (iv) for diametrical x and y it follows that $z = x$ and $\bar{z} = y$. Consider the set of all possible sequences of points connecting the diametrical points x_1 and x_4 and consisting of $\binom{3}{2} + 1 = 4$ points:

$$\begin{aligned} (a_1) & x_1, x_6, x_5, x_4; & (a_2) & x_1, x_6, x_2, x_4; & (a_3) & x_1, x_6, x_3, x_4; \\ (a_4) & x_1, x_5, x_2, x_4; & (a_5) & x_1, x_5, x_3, x_4; & (a_6) & x_1, x_5, x_6, x_4. \end{aligned}$$

Without loss of generality, the symmetric case where the points x_2 or x_3 are at the second place is omitted. Sequences (a_3) and (a_4) can not satisfy (iv) since they contain consecutive diametrical pairs of points. Let us consider the sequence (a_1) . Suppose condition (iv) holds for this case, i.e.,

$$a + b + c = \text{diam}X.$$

By (12) we have

$$b + c = d, \quad a + c = e, \quad a + b = f.$$

In this case (X, d) is well-defined, i.e., all triangle inequalities are satisfied. Using (2) and Fig. 2 we see that (X, d) is isometric to (S_3, d_W) with the isometry $\Phi: X \rightarrow S_3$:

$$\begin{aligned} \Phi(x_1) &= (1, 2, 3), & \Phi(x_2) &= (2, 1, 3), & \Phi(x_3) &= (3, 1, 2), \\ \Phi(x_4) &= (3, 2, 1), & \Phi(x_5) &= (2, 3, 1), & \Phi(x_6) &= (1, 3, 2), \end{aligned}$$

and the weight

$$W = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

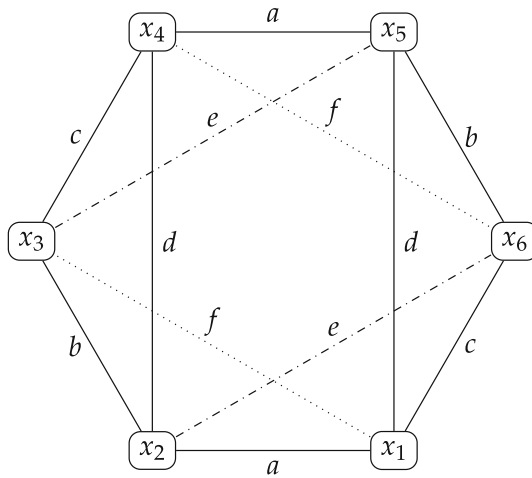


FIGURE 3. The metric space (X, d) with $X = \{x_1, \dots, x_6\}$

Note that this isometry is not necessarily unique.

Consider case (a_2) . Again, suppose that condition (iv) holds, i.e.,

$$c + d + e = \text{diam}X.$$

Using (12) we have

$$c + e = a, \quad c + d = b, \quad d + e = f.$$

In this case (X, d) is isometric to (S_3, d_W) , for example, with the isometry $\Phi: X \rightarrow S_3$:

$$\begin{aligned} \Phi(x_1) &= (1, 2, 3), & \Phi(x_6) &= (1, 3, 2), & \Phi(x_2) &= (2, 3, 1), \\ \Phi(x_4) &= (3, 2, 1), & \Phi(x_3) &= (3, 1, 2), & \Phi(x_5) &= (2, 1, 3), \end{aligned}$$

and the weight

$$W = \begin{pmatrix} 0 & d & e \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

Cases (a_5) , (a_6) are analogous.

Since by the statement of conjecture condition (iv) holds for (X, d) , it holds at least for one of the cases (a_1) , (a_2) , (a_5) , (a_6) or for their respective symmetric cases which were omitted from consideration. The existence of isometry in each case is shown. \square

5. Conclusion

Usually, the concept of a metric is associated with the distance between points of a certain space. But in mathematics there are a lot of metrics defined not on points but on completely different mathematical objects. A large number of distances is collected in [7]. Among these distances one can distinguish distances on graphs, matrices, strings, etc. In this work we have considered a metric space the points of which are permutations S_n of the numbers $1, \dots, n$ with fixed n . The introduced metric d_W generalizes not only the well-known Kendall τ metric but also some another its weighted generalizations.

The paper is devoted to the study of geometric properties of the space (S_n, d_W) . It is proved that (S_n, d_W) in general case is a pseudometric space and is a metric space if and only if all weights in the strictly upper triangular matrix W are positive. Some basic geometric properties of this space are also described. The observation that the vertex set of a permutohedron of order n coincides with the set of points of the space (S_n, d_W) allowed us to see that the edge-graph G_n of such permutohedron can be used as a convenient tool for studying the space (S_n, d_W) . Using the graph G_n we give a criterion which guarantees that some point “lies between” another two fixed points from (S_n, d_W) and describe special type four-point subsets of (S_n, d_W) so called “pseudolinear quadruples”. At the end we formulate a conjecture that characterizes the metric space (S_n, d_W) and prove it for the case $n = 3$.

Acknowledgements

The authors thank the anonymous referees for their remarks which considerably improved this article.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [1] Aboulker, P., Chen, X., Huzhang, G., Kapadia, R., Supko, C.: Lines, betweenness and metric spaces. *Discrete Comput. Geom.* **56**(2), 427–448 (2016)
- [2] Brandenburg, F.J., Gleißner, A., Hofmeier, A.: Comparing and aggregating partial orders with Kendall tau distances. *Discrete Math. Algorithms Appl.* **5**(2), 88–99 (2013)
- [3] Buzaglo, S., Etzion, T.: Perfect permutation codes with the Kendall's τ -metric. In: 2014 IEEE International Symposium on Information Theory, pp. 2391–2395. IEEE, Honolulu, HI, USA (2014)
- [4] Chee, Y.M., Vu, V.K.: Breakpoint analysis and permutation codes in generalized Kendall tau and Cayley metrics. In: 2014 IEEE International Symposium on Information Theory, pp. 2959–2963. IEEE, Honolulu, HI, USA (2014)
- [5] Cicirello, V.A.: Kendall tau sequence distance: Extending Kendall tau from ranks to sequences. *EAI Endorsed Trans. Ind. Netw. and Intell. Syst.* **7**(23), 1–12 (2020)
- [6] Deza, M., Huang, T.: Metrics on permutations, a survey. *J. Comb. Inf. Syst. Sci.* **23**, 173–185 (1998)
- [7] Deza, M.M., Deza, E.: *Encyclopedia of distances*, fourth edn. Springer, Berlin (2016)
- [8] Dovgoshei, A., Petrov, E.: Ptolemaic spaces. *Siberian Math. J.* **52**(2), 222–229 (2011)
- [9] Farnoud, F., Touri, B., Milenkovic, O.: Novel distance measures for vote aggregation. arXiv preprint [arXiv:1203.6371](https://arxiv.org/abs/1203.6371) (2012)
- [10] Fischer, S., Schumann, A., Schnurr, A.: Ordinal pattern dependence between hydrological time series. *J. Hydrol.* **548**, 536–551 (2017)
- [11] Keller, K., Wittfeld, K.: Distances of time series components by means of symbolic dynamics. *Internat. J. Bifur. Chaos* **14**, 693–703 (2004)
- [12] Kendall, M.: A new measure of rank correlation. *Biometrika* **30**(1–2), 81–89 (1938)
- [13] Kruskal, W.H.: Ordinal measures of association. *J. Amer. Statist. Assoc.* **53**(284), 814–861 (1958)

- [14] Kumar, R., Vassilvitskii, S.: Generalized distances between rankings. In: Proceedings of the 19th International Conference on World Wide Web, WWW '10, pp. 571–580. ACM, New York, NY, USA (2010)
- [15] Lee, P.H., Yu, P.L.H.: Distance-based tree models for ranking data. *Comput. Statist. Data Anal.* **54**(6), 1672–1682 (2010)
- [16] Lee, P.H., Yu, P.L.H.: Mixtures of weighted distance-based models for ranking data with applications in political studies. *Comput. Statist. Data Anal.* **56**(8), 2486–2500 (2012)
- [17] Liu, C., Han, J.: Failure proximity: a fault localization-based approach. In: Proceedings of the 14th ACM SIGSOFT international symposium on Foundations of software engineering, pp. 46–56. Association for Computing Machinery, New York, NY, USA (2006)
- [18] Liu, C., Zhang, X., Han, J.: A systematic study of failure proximity. *IEEE Trans. on Softw. Eng.* **34**(6), 826–843 (2008)
- [19] Menger, K.: Untersuchungen über allgemeine Metrik. *Math. Ann.* **100**(1), 75–163 (1928)
- [20] Ouyang, G., Dang, C., Richards, D.A., Li, X.: Ordinal pattern based similarity analysis for eeg recordings. *Clin. Neurophysiol.* **121**(5), 694–703 (2010)
- [21] Shieh, G.S.: A weighted Kendall's tau statistic. *Statist. Probab. Lett.* **39**(1), 17–24 (1998)
- [22] Spearman, C.: The proof and measurement of association between two things. *Amer. J. of Psych.* **15**(1), 72–101 (1904)
- [23] Venkatraghavan, V., Bron, E.E., Niessen, W.J., Klein, S.: A discriminative event based model for Alzheimer's disease progression modeling. In: Information Processing in Medical Imaging, pp. 121–133. Springer International Publishing, Cham (2017)
- [24] Wang, J., Shang, P., Shi, W., Cui, X.: Dissimilarity measure based on ordinal pattern for physiological signals. *Commun. Nonlinear Sci. Numer. Simul.* **37**, 115–124 (2016)
- [25] Ziegler, G.: Lectures on polytopes. Graduate Texts in Mathematics, 152. Springer-Verlag, New York (1995)

Albert Bruno Piek
Institute of Mathematics
University of Lübeck
Ratzeburger Allee 160
23562 Lübeck
Germany
e-mail: piek@math.uni-luebeck.de

Evgeniy Petrov
Institute of Applied Mathematics and Mechanics of the NAS of Ukraine
Dobrovolskogo Str. 1
84100 Slovyansk
Ukraine
e-mail: eugeniy.petrov@gmail.com

Communicated by Kolja Knauer

Received: 9 February 2020.

Accepted: 12 December 2020.