

**DIGRAPHS OF POTENTIAL BARRIERS:
PROPERTIES OF THEIR TREE STRUCTURE
AND ALGORITHM FOR CONSTRUCTING
MINIMUM SPANNING FORESTS**V.A. BUSLOV 

Abstract: For a weighted digraph without loops V , the arc weights of which can be obtained from an undirected graph with loops P according to the rule $v_{ij} = p_{ij} - p_{ii}$, the properties are studied. An effective algorithm for constructing directed trees of minimum weight and an algorithm for constructing spanning directed forests of minimum weight are proposed.

Keywords: weighted digraph, spanning forest, minimum tree.

BUSLOV, V.A., DIGRAPHS OF POTENTIAL BARRIERS: PROPERTIES OF THEIR TREE STRUCTURE AND ALGORITHM FOR CONSTRUCTING MINIMUM SPANNING FORESTS.

© 202X BUSLOV V.A..

Received November, X, 202X, Published December, XX, 202X.

Introduction

This article is based on the results of works [10]-[12]. The notations and definitions correspond to those adopted in these two works, up to some fonts and specially specified simplifications.

1 Basic notations and definitions

The work uses both directed and undirected graphs, forests and trees. For brevity, we use the terms graph, forest and tree, denoting both undirected and directed, if this does not lead to misunderstandings. Where necessary, we indicate exactly what type of graphs we have in mind. In the first part of the work only digraphs appear.

For a digraph G , we denote the set of its vertices by VG , and the set of arcs by AG .

We consider spanning subgraphs (with a set of vertices N) that are entering (incoming) forests. An entering forest is a digraph in which no more than one arc comes from each vertex and there are no contours. A tree is a connected component of a forest. The root of a tree (forest) is the vertex from which arc does not come. Let T_i^F denote the tree of forest F rooted at vertex i . We denote the set of roots of the forest F by K_F . The outgoing forest is obtained by replacing all the arcs of the entering forest with inverse ones (then the root is the vertex into which arcs do not go).

Of the directed ones, only entering forests are used, in the future simply — forests, until the moment when undirected graphs do not appear in the presentation. An undirected forest is a graph without cycles, its connected components are trees.

The subgraph H of a graph G induced by the set $S \subset VG$ (or the restriction of the graph G to the set S) is a subgraph in which $VH = S$, and the set of its arcs is all arcs of the graph G , both ends of which belong to the set S . For it we use the notation $H = G|_S$.

If there is an arc whose outcome belongs to the set S , but its entry does not, then we say that the arc comes from the set S . Similarly, if there is an arc whose entry belongs to S and whose outcome does not belong, then we say that the arc enters S .

The outgoing neighborhood $N_S^{out}(G)$ of the set S is the set of entries of arcs coming in graph G from the set S ; The incoming neighborhood $N_S^{in}(G)$ is defined similarly.

For a subgraph G of a graph V and a subset of the vertex set $\mathbf{S} \subseteq \mathbf{N}$, we introduce weights

$$\gamma_{\mathbf{S}}^G = \sum_{\substack{i \in \mathbf{S} \\ (i,j) \in AG}} v_{ij}, \quad \gamma^G = \gamma_{\mathbf{N}}^G = \sum_{(i,j) \in AG} v_{ij}. \quad (1)$$

The value $\gamma_{\mathbf{S}}^G$ is also formed from arcs whose entries do not belong to the set \mathbf{S} . If in a graph G no arcs come from the set \mathbf{S} itself, then $\gamma_{\mathbf{S}}^G = \gamma^{G|\mathbf{S}}$.

\mathcal{F}^k is a set of forests consisting of $k = 1, 2, \dots, N$ trees. We denote the minimum weight γ^F among all forests $F \in \mathcal{F}^k$ by φ^k :

$$\varphi^k = \min_{F \in \mathcal{F}^k} \gamma^F. \quad (2)$$

If $\mathcal{F}^k = \emptyset$, we set $\varphi^k = \infty$, in particular, $\varphi^0 = \infty$. The set \mathcal{F}^N consists of only one empty forest and $\varphi^N = 0$. Any forest in \mathcal{F}^{N-1} contains exactly one arc, so $\varphi^{N-1} = \min_{(i,j) \in AV} v_{ij}$.

$\tilde{\mathcal{F}}^k$ is a subset of the set of forests \mathcal{F}^k on which the minimum of φ^k is achieved: $F \in \tilde{\mathcal{F}}^k \Leftrightarrow F \in \mathcal{F}^k$ and $\gamma^F = \varphi^k$. Forests from $\tilde{\mathcal{F}}^k$ are called minimal.

$\mathcal{F}^k|_{\mathbf{S}}$ — the set of subgraphs of k -component forests induced by the set \mathbf{S} ;

$\tilde{\mathcal{F}}^k|_{\mathbf{S}}$ — the set of subgraphs of k -component minimum weight forests induced by the set \mathbf{S} ;

$F_{\uparrow \mathbf{S}}^G$ — the graph obtained from F by replacing the arcs coming from the vertices of the set \mathbf{S} with arcs coming from the same vertices in graph G .

For any subset $\mathbf{S} \subset \mathbf{N}$, its complement $\bar{\mathbf{S}} = \mathbf{N} \setminus \mathbf{S}$.

Forest $F \in \mathcal{F}^k$ is called a *pseudo-ancestor* of a forest $G \in \mathcal{F}^{k-1}$, and G is a *pseudo-descendant* of F , if there is a root j of F such that: $\mathbf{K}_G = \mathbf{K}_F \setminus \{j\}$ and $G|_{VT_q^F} = T_q^F$ for $q \in \mathbf{K}_G$. Such forests are called pseudo-related, and the set of such pseudo-descendants of F is denoted by \mathcal{P}_j^F .

Forest $F \in \mathcal{F}^k$ is called an *ancestor* of a forest $G \in \mathcal{F}^{k-1}$, and G is called a *descendant* of a forest F , if for some $j \in \mathbf{K}_F$ the following holds: $G \in \mathcal{P}_j^F$ and $G|_{VT_j^F}$ is a tree. Forests F and G are called related, and the set of such descendants of a forest F is denoted by \mathcal{R}_j^F . The set of all descendants of a forest F is denoted by $\mathcal{R}^F = \bigcup_{j \in \mathbf{K}_F} \mathcal{R}_j^F$.

2 Properties used

2.1. Arc replacement operation. We will need a simple property [10, Corollary 2 of Lemma 1] of the operation of replacing arcs ing from the vertices of a certain set in two arbitrary forests. Let us formulate it in a somewhat broader setting.

Property 1. [10] *Let F and G be forests, $\forall G \subseteq \forall F$, let $D \subseteq \forall G$ be a subset of the vertex set that is not entered by arcs in the forest F . Then graph $F_{\uparrow D}^G$ is a forest.*

A thinner one will also be used.

Property 2. [10, Corollary 6 from Lemma 1] *Let F and G be forests with the same vertex set, and let T^F and T^G be trees of the forests of F and G , respectively, and let $D \subset \forall T^F \cap \forall T^G$, such that $N_D^{in}(G) = \emptyset$ and $N_D^{out}(G) \subset \forall T^G \setminus \forall T^F$. Then the graphs $F_{\uparrow D}^G$ and $G_{\uparrow D}^F$ are forests.*

2.2. Minima of weight on subsets of a vertex set. In [12], for any subset S of the set of all vertices N , special tree-like minima are defined:

$$\lambda_S^{\bullet q} = \min_{T \in \mathcal{T}_S^{\bullet q}} \gamma^T, \quad \lambda_S^\bullet = \min_{q \in S} \lambda_S^{\bullet q}, \quad (3)$$

where $\mathcal{T}_S^{\bullet q}$ is a set of trees with vertex set S and root at vertex $q \in S$. The second of the minima (3) can be rewritten as

$$\lambda_S^\bullet = \min_{T \in \mathcal{T}_S^\bullet} \gamma^T, \quad \mathcal{T}_S^\bullet = \bigcup_{q \in S} \mathcal{T}_S^{\bullet q}, \quad (4)$$

using the disjoint union of sets $\mathcal{T}_S^{\bullet q}$ by vertices $q \in S$.

Let us also identify subsets of trees on which the corresponding minima are achieved: $T \in \tilde{\mathcal{T}}_S^{\bullet q}$ if $T \in \mathcal{T}_S^{\bullet q}$ and $\gamma^T = \lambda_S^{\bullet q}$; $T \in \tilde{\mathcal{T}}_S^\bullet$ if $T \in \mathcal{T}_S^\bullet$ and $\gamma^T = \lambda_S^\bullet$.

Also in [12] forest-like minima are defined

$$\mu_S^{\bullet q} = \min_{F \in \mathcal{F}_S^{\bullet q}} \gamma_S^F, \quad \mu_S^\bullet = \min_{q \in S} \mu_S^{\bullet q}, \quad (5)$$

where $\mathcal{F}_S^{\bullet q}$ is the set of spanning forests such that if $F \in \mathcal{F}_S^{\bullet q}$, then the set S contains exactly one root of forest F , namely vertex $q \in S$. In this case, the entries of arcs coming in the forest F from the vertices of the set S may or may not belong to the set S itself. Arcs coming in F from the vertices of \bar{S} do not affect the value $\mu_S^{\bullet q}$.

For these quantities the following is true

Property 3. [12, Proposition 2] *Let $F \in \tilde{\mathcal{F}}^k$ and $l \in K_F$, then*

$$\mu_{\mathbf{S}}^{\bullet l} = \mu_{\mathbf{S}}^{\bullet} = \lambda_{\mathbf{S}}^{\bullet} = \lambda_{\mathbf{S}}^{\bullet l} = \Upsilon_{\mathbf{S}}^F, \quad (6)$$

where $\mathbf{S} = \mathbf{V}T_l^F$.

Forest-like minima are also defined for the case when the set \mathbf{S} does not contain forest roots:

$$\mu_{\mathbf{S}}^{\circ} = \min_{F \in \mathcal{F}_{\mathbf{S}}^{\circ}} \Upsilon_{\mathbf{S}}^F, \quad (7)$$

where $\mathcal{F}_{\mathbf{S}}^{\circ}$ is a set of spanning forests in which arcs come from all vertices of the set \mathbf{S} . Arcs come from the vertices of $\bar{\mathbf{S}}$ in any forest $F \in \mathcal{F}_{\mathbf{S}}^{\circ}$ are arbitrary and do not affect the weight value.

For all introduced weights, both tree-like (type λ) and forest-like (type μ), we assume that they are equal to infinity if the corresponding set of trees (\mathcal{T}) and, accordingly, forests (\mathcal{F}) is empty.

3 Tree-like weights on a subset not containing roots

We are further interested in the situation when arcs come from all vertices of the set \mathbf{S} and they form a special tree. Namely, we introduce a set of trees $\mathcal{T}_{\mathbf{S}}^{\circ q}$, which are subgraphs of the original graph V , and if $T \in \mathcal{T}_{\mathbf{S}}^{\circ q}$, then $|\mathbf{V}T| = |\mathbf{S}| + 1$ and $T|_{\mathbf{S}} \in \mathcal{T}_{\mathbf{S}}^{\bullet q}$. This means that in the tree T , from the set \mathbf{S} itself, there is exactly one arc, and it comes from the vertex q . Let this be the arc (q, r) . Obviously, $r \in \bar{\mathbf{S}}$ and this vertex is the root of the tree T . Despite the fact that r is the root, this vertex is not of particular interest in the context of this work. Therefore, it is not an index in the set of trees under consideration. But the vertex q is more significant, since the subject of consideration is the arcs coming from the vertices of the set \mathbf{S} , to which the vertex q belongs, unlike r . Note that the weights of the trees of the corresponding trees T and $T_q = T|_{\mathbf{S}}$ are related by the ratio

$$\Upsilon^T = \Upsilon^{T_q} + v_{qr}. \quad (8)$$

Let us also introduce the disjoint union

$$\mathcal{T}_{\mathbf{S}}^{\circ} = \bigcup_{q \in \mathbf{S}} \mathcal{T}_{\mathbf{S}}^{\circ q} \quad (9)$$

and determine the weights

$$\lambda_{\mathbf{S}}^{\circ} = \min_{T \in \mathcal{T}_{\mathbf{S}}^{\circ}} \Upsilon^T, \quad \lambda_{\mathbf{S}}^{\circ q} = \min_{T \in \mathcal{T}_{\mathbf{S}}^{\circ q}} \Upsilon^T. \quad (10)$$

Obviously

$$\lambda_{\mathbf{S}}^{\circ} = \min_{q \in \mathbf{S}} \lambda_{\mathbf{S}}^{\circ q} . \quad (11)$$

Note that for different trees T and T' from $\mathcal{T}_{\mathbf{S}}^{\circ}$, the sets of their vertices, generally speaking, do not coincide, since the root of the tree can be any vertex from $\bar{\mathbf{S}}$.

Let us find out the connections between the introduced weights. If the tree $T \in \mathcal{T}_{\mathbf{S}}^{\circ}$ is supplemented with the remaining vertices to a spanning subgraph, then we obtain a spanning forest belonging to the set $\mathcal{F}_{\mathbf{S}}^{\circ}$. Therefore, the following inequality is satisfied

$$\mu_{\mathbf{S}}^{\circ} \leq \lambda_{\mathbf{S}}^{\circ} . \quad (12)$$

Note that for an arbitrary \mathbf{S} , [12] a similar inequality $\mu_{\mathbf{S}}^{\bullet} \leq \lambda_{\mathbf{S}}^{\bullet}$ holds. Equality (6) holds if \mathbf{S} is the vertex set of some tree in the minimal forest. A similar situation holds for inequality (12).

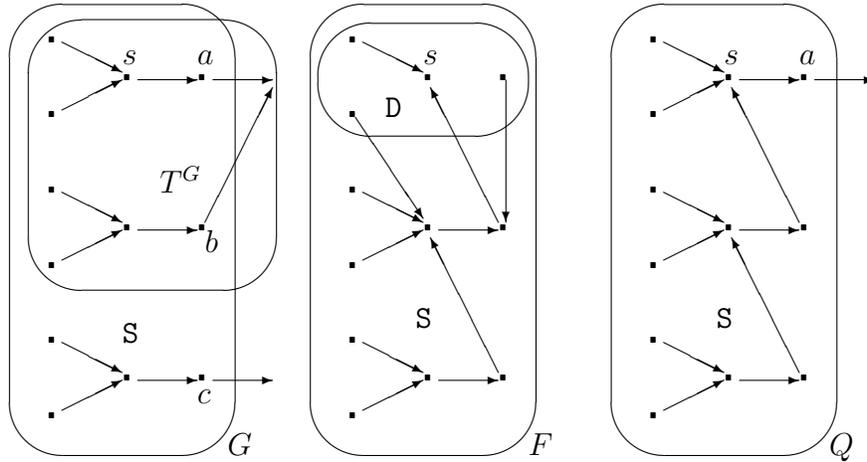


FIG. 1. On the left are the arcs of the forest G coming from the vertices of the set $\mathbf{S} = \mathbf{V}T_s^F$; in the center are the arcs of the tree T_s^F and the set \mathbf{D} of vertices of the connected component of the induced subgraph $G|_{\mathbf{S}}$, including the vertex s ; on the right are the arcs of the graph $Q = F_{\uparrow \mathbf{D}}^G$ coming from the vertices of the set \mathbf{S} .

Theorem 1. *Let $F \in \tilde{\mathcal{F}}^k$ and $s \in \mathbf{K}_F$, then*

$$\mu_{\mathbf{S}}^{\circ} = \lambda_{\mathbf{S}}^{\circ}, \quad (13)$$

where $\mathbf{S} = \mathbf{V}T_s^F$.

Proof. First of all, we note that if $\mu_{\mathbf{S}}^{\circ} = \infty$, then by (12) $\lambda_{\mathbf{S}}^{\circ} = \infty$. Therefore, we consider only the situation when $\mu_{\mathbf{S}}^{\circ} < \infty$. Let the spanning forest $G \in \mathcal{F}_{\mathbf{S}}^{\circ}$ be such that $\Upsilon_{\mathbf{S}}^G = \mu_{\mathbf{S}}^{\circ}$. For convenience and without loss of generality, we assume that no arcs come from the vertices of the set $\bar{\mathbf{S}}$ ($\Upsilon^G = \Upsilon_{\mathbf{S}}^G$).

The induced subgraph $G|_{\mathbf{S}}$ is a forest on the vertex set \mathbf{S} . The roots of its trees are those vertices of the set \mathbf{S} from which arcs in the graph G itself originate and enter $\bar{\mathbf{S}}$ (in Fig. 1 these are vertices a , b , and c). Let \mathbf{D} be the set of vertices of the tree of the induced forest $G|_{\mathbf{S}}$ that contains the root s of the forest F , and let T^G be the tree of the forest G that contains \mathbf{D} .

Let us introduce graphs $Q = F_{\uparrow \mathbf{D}}^G$ and $R = G_{\uparrow \mathbf{D}}^F$. The set \mathbf{D} satisfies the conditions of Property 2, so both graphs Q and R are forests. Note that $Q \in \mathcal{R}_s^F$.

In the forest Q , arcs come from all vertices of the set \mathbf{S} , and exactly one arc comes from the set \mathbf{S} itself (in Fig. 1 on the right, it comes from vertex a). Thus, $Q|_{\mathbf{S}}$ is a tree with a root at a . If we supplement it with an arc coming from a in the forest Q , we obtain a tree belonging to the set $\mathcal{T}_{\mathbf{S}}^{\circ}$. Thus $\Upsilon_{\mathbf{S}}^Q \geq \lambda_{\mathbf{S}}^{\circ}$. We have

$$\lambda_{\mathbf{S}}^{\circ} \leq \Upsilon_{\mathbf{S}}^Q = \Upsilon_{\mathbf{S}}^F - \Upsilon_{\mathbf{D}}^F + \Upsilon_{\mathbf{D}}^G = \lambda_{\mathbf{S}}^{\bullet} - \Upsilon_{\mathbf{D}}^F + \Upsilon_{\mathbf{D}}^G. \quad (14)$$

In the forest R arcs originate from all the vertices of the set \mathbf{S} except one (this is the vertex s), so $\Upsilon_{\mathbf{S}}^R \geq \mu_{\mathbf{S}}^{\bullet}$ is executed. On the other hand, considering (6)

$$\lambda_{\mathbf{S}}^{\bullet} = \mu_{\mathbf{S}}^{\bullet} \leq \Upsilon_{\mathbf{S}}^R = \Upsilon_{\mathbf{S}}^G - \Upsilon_{\mathbf{D}}^G + \Upsilon_{\mathbf{D}}^F = \mu_{\mathbf{S}}^{\circ} - \Upsilon_{\mathbf{D}}^G + \Upsilon_{\mathbf{D}}^F. \quad (15)$$

Substituting $\lambda_{\mathbf{S}}^{\bullet}$ from (15) into (14), we obtain $\lambda_{\mathbf{S}}^{\circ} \leq \mu_{\mathbf{S}}^{\circ}$, and by (12) the reverse inequality holds. \square

Proposition 1. *Let $\mathbf{S} \subsetneq \mathbb{N}$ and $\mathcal{T}_{\mathbf{S}}^{\circ} \neq \emptyset$, then*

$$\lambda_{\mathbf{S}}^{\circ} = \min_{q \in \mathbf{S}} \left(\lambda_{\mathbf{S}}^{\bullet q} + \min_{r \notin \mathbf{S}} v_{qr} \right). \quad (16)$$

Proof. Let $q \in \mathbf{S}$. By definition, if $T \in \mathcal{T}_{\mathbf{S}}^{\circ q}$, then its subgraph $T|_{\mathbf{S}}$ induced by the set \mathbf{S} is a tree. Let $T_q = T|_{\mathbf{S}}$. Then

$$\Upsilon^T = \Upsilon^{T_q} + v_{qr}, \quad (17)$$

where $(q, r) \in AT$. Since T_q is a tree, this arc is the only one coming from q at T . Moreover, $r \notin \mathbf{S}$ is the root of the tree T , and $T_q \in \mathcal{T}_{\mathbf{S}}^{\bullet q}$. Let us determine the minimum of (17) for a fixed q .

$$\lambda_{\mathbf{S}}^{\circ q} = \min_{T \in \mathcal{T}_{\mathbf{S}}^{\circ q}} \Upsilon^T = \min_{\substack{r \notin \mathbf{S} \\ T_q \in \mathcal{T}_{\mathbf{S}}^{\bullet q}}} (\Upsilon^{T_q} + v_{qr}). \quad (18)$$

Since $r \notin \mathbf{S}$ and this vertex is not related to the weight of the tree T_q , the expression for the minimum takes the form

$$\lambda_{\mathbf{S}}^{\circ q} = \min_{T_q \in \mathcal{T}_{\mathbf{S}}^{\bullet q}} \Upsilon^{T_q} + \min_{r \notin \mathbf{S}} v_{qr} .$$

Given the definition (3), we obtain

$$\lambda_{\mathbf{S}}^{\circ q} = \lambda_{\mathbf{S}}^{\bullet q} + \min_{r \notin \mathbf{S}} v_{qr} . \quad (19)$$

In turn, (11) now implies (16). \square

4 Relatedness of forests

The main theorem [10, Theorem 2 (on related forests)] we formulate as a property

Property 4. *Let $\mathcal{F}^{k-1} \neq \emptyset$, then any forest from $\tilde{\mathcal{F}}^k$ has a descendant in $\tilde{\mathcal{F}}^{k-1}$ and vice versa – any forest from $\tilde{\mathcal{F}}^{k-1}$ has an ancestor in $\tilde{\mathcal{F}}^k$.*

In [12, Theorem 2] a criterion is formulated for a pseudo-descendant of a minimal forest to also be minimal, and on this basis an efficient algorithm for constructing minimal pseudo-sibling forests is presented. We present the formulation in the form of a property.

Property 5. *Let $F \in \tilde{\mathcal{F}}^k$, $j \in \mathbf{K}_F$ and $G \in \mathcal{P}_j^F$. Then in order for $G \in \tilde{\mathcal{F}}^{k-1}$ it is necessary and sufficient that*

$$\Upsilon_{\mathbf{v}_{T_j^F}}^G = \mu_{\mathbf{v}_{T_j^F}}^{\circ} \quad (20)$$

and

$$\mu_{\mathbf{v}_{T_j^F}}^{\circ} - \mu_{\mathbf{v}_{T_j^F}}^{\bullet} = \min_{l \in \mathbf{K}_F} \left(\mu_{\mathbf{v}_{T_l^F}}^{\circ} - \mu_{\mathbf{v}_{T_l^F}}^{\bullet} \right) . \quad (21)$$

Let us prove a similar criterion for related forests.

Theorem 2. *Let $F \in \tilde{\mathcal{F}}^k$, $y \in \mathbf{K}_F$ and $G \in \mathcal{R}_y^F$. In order for $G \in \tilde{\mathcal{F}}^{k-1}$, it is necessary and sufficient that*

$$\Upsilon_{\mathbf{v}_{T_y^F}}^G = \lambda_{\mathbf{v}_{T_y^F}}^{\circ} \quad (22)$$

and

$$\lambda_{\mathbf{v}_{T_y^F}}^{\circ} - \lambda_{\mathbf{v}_{T_y^F}}^{\bullet} = \min_{l \in \mathbf{K}_F} \left(\lambda_{\mathbf{v}_{T_l^F}}^{\circ} - \lambda_{\mathbf{v}_{T_l^F}}^{\bullet} \right) . \quad (23)$$

Proof. By Property 3 and Theorem 1, for all trees in forest F ,

$$\lambda_{\mathbf{v}_{T_l^F}}^{\bullet} = \mu_{\mathbf{v}_{T_l^F}}^{\bullet} , \quad \lambda_{\mathbf{v}_{T_l^F}}^{\circ} = \mu_{\mathbf{v}_{T_l^F}}^{\circ} , \quad l \in \mathbf{K}_F . \quad (24)$$

Forest G is a descendant of forest F , and in particular is a pseudo-descendant. If forest G and root y satisfy conditions (22) and (23),

then forest G also satisfies conditions (20) and (21). Thus, by Property 5, forest G is a minimal pseudo-descendant of forest F . The weights of all minimal forests with a fixed number of roots coincide. So the forest G is a minimal descendant. In the opposite direction. Let G be a minimal descendant and $G \in \mathcal{R}_j^F$. Then it is a minimal pseudo-descendant. By Property 4, conditions (20) and (21) are satisfied for it. That is, conditions (22) and (23). \square

Corollary 1. *Let $F \in \tilde{\mathcal{F}}^k$, then*

$$\varphi^{k-1} - \varphi^k = \min_{l \in \mathbb{K}_F} \left(\lambda_{\mathbb{V}T_l^F}^\circ - \lambda_{\mathbb{V}T_l^F}^\bullet \right) . \quad (25)$$

Proof. Let $G \in \tilde{\mathcal{F}}^{k-1}$ be a descendant of forest F and $\{y\} = \mathbb{K}_F \setminus \mathbb{K}_G$. That is, y is the root on which the minimum on the right-hand side of (25) is reached, that is, (23) is satisfied. We have: $\varphi^k = \Upsilon^F$, $\varphi^{k-1} = \Upsilon^G$. Forests F and G differ only in the arcs coming from the vertices of the set $\mathbb{V}T_y^F$. Thus, taking into account (22)

$$\varphi^{k-1} - \varphi^k = \Upsilon^G - \Upsilon^F = \Upsilon_{\mathbb{V}T_y^F}^G - \Upsilon_{\mathbb{V}T_y^F}^F = \lambda_{\mathbb{V}T_y^F}^\circ - \lambda_{\mathbb{V}T_y^F}^\bullet . \quad (26)$$

Taking into account (23), we obtain (25). \square

5 Recalculation of minimum weights when enlarging trees

Proposition 2. *Let $F \in \tilde{\mathcal{F}}^k$, $y \in \mathbb{K}_F$, $G \in \tilde{\mathcal{F}}^{k-1} \cap \mathcal{R}_y^F$. Let also the entry of the only arc coming from the set $\mathbb{V}T_y^F$ at G belong to the set $\mathbb{V}T_x^F$. Then*

$$\lambda_{\mathbb{V}T_x^G}^\bullet = \lambda_{\mathbb{V}T_x^F}^\bullet + \lambda_{\mathbb{V}T_y^F}^\circ . \quad (27)$$

Proof. First of all, we note that in the forest G all trees except the tree rooted at x are the same as in the forest F and nothing has changed for the weight minima on their vertex sets. Since the forest G itself is minimal, then by Property 3 for the vertex set \mathbb{S} of any tree in it, $\lambda_{\mathbb{S}}^\bullet = \Upsilon_{\mathbb{S}}^G$ holds. The only tree in G with a vertex set different from that in F is the tree T_x^G . Since $G \in \mathcal{R}_y^F$, then $\mathbb{V}T_x^G = \mathbb{V}T_x^F \cup \mathbb{V}T_y^F$ and $G|_{\mathbb{V}T_x^F} = \mathbb{V}T_x^F$. Taking into account (20), we have:

$$\lambda_{\mathbb{V}T_x^G}^\bullet = \Upsilon_{\mathbb{V}T_x^G}^G = \Upsilon_{\mathbb{V}T_x^F}^G + \Upsilon_{\mathbb{V}T_y^F}^G = \Upsilon_{\mathbb{V}T_x^F}^F + \Upsilon_{\mathbb{V}T_y^F}^G = \lambda_{\mathbb{V}T_x^F}^\bullet + \lambda_{\mathbb{V}T_y^F}^\circ . \quad \square$$

Propositions 1 and 2 together with Theorem 2 allow us to algorithmically construct related minimal forests. However, direct use of (16) in finding the values of $\lambda_{\mathfrak{S}}^{\circ}$ (and the corresponding trees) is not effective, since it is required to sequentially assign each vertex q of the set $\mathfrak{S} = VT_y^F$ as a root. Property 5 allows us to avoid this procedure and define more convenient for calculation minima $\mu_{\mathfrak{S}}^{\circ}$ and the corresponding forests [12]. However, there is an important, especially in the physics of Brownian motion in a potential drift field ([13], [14]), special case of weight function, when it is more efficient to calculate the values $\lambda_{\mathfrak{S}}^{\circ}$ and find the trees corresponding to them.

6 Digraph of (potential) barriers and graph of potential

6.1. Clarification of the use of terminology. Further, both directed and undirected graphs will be used, which means that each time it is necessary to stipulate whether an directed graph or a graph (undirected) is used. Therefore, now the widespread use of the term graph (as well as the terms forest and tree) becomes inconvenient and can lead to confusion. The term graph itself now implies an undirected graph. Everything that was previously called forests and trees, in the strict sense, are directed forests and directed trees, and entering ones at that. Unless specifically noted or clear from the context, the terms forest and tree now refer to undirected graphs. Namely, a forest is an acyclic graph, and a connected component of a forest is a tree.

6.2. Notations and definitions for undirected graphs. To denote undirected graphs we use fonts without italics. Along with the original weighted directed graph V , there will be an associated undirected graph P with loops on the same set of vertices: $VP = N$. The set of edges (and loops) is denoted by EP . Depending on the context, the pair (i, j) , for $i \neq j$, can be considered both as an arc of the directed graph V and as an edge of the graph P . For the graph P , the pairs (i, j) and (j, i) define the same edge with weight p_{ij} .

For a subgraph G of a graph P and a subset of the vertex set $\mathfrak{S} \subseteq N$, we introduce weights

$$\Upsilon_{\mathfrak{S}}^G = \sum_{\substack{i, j \in \mathfrak{S} \\ (i, j) \in EG}} p_{ij}, \quad \Upsilon^G = \Upsilon_N^G = \sum_{(i, j) \in EG} p_{ij}. \quad (28)$$

Here, in contrast to the oriented situation (1), always $\Upsilon_{\mathfrak{S}}^G = \Upsilon^{G|\mathfrak{S}}$, but there is no additivity property.

The sets of spanning forests of the graph P consisting of k trees are denoted by F^k . Forests F from F^k on which the minimum of weight Υ^F is achieved are called minimal. The set of such forests is denoted by \tilde{F}^k .

Similarly, if $S \subseteq \mathbb{N}$, then by T_S we mean the set of trees T that are subgraphs of P such that $VT = S$. Trees $T \in T_S$ on which the minimum of weight Υ^T is achieved are called minimal on the set S . The set of such trees is denoted by \tilde{T}_S , and the minimum of weight itself is denoted by ν_S :

$$\nu_S = \min_{T \in \tilde{T}_S} \Upsilon^T . \quad (29)$$

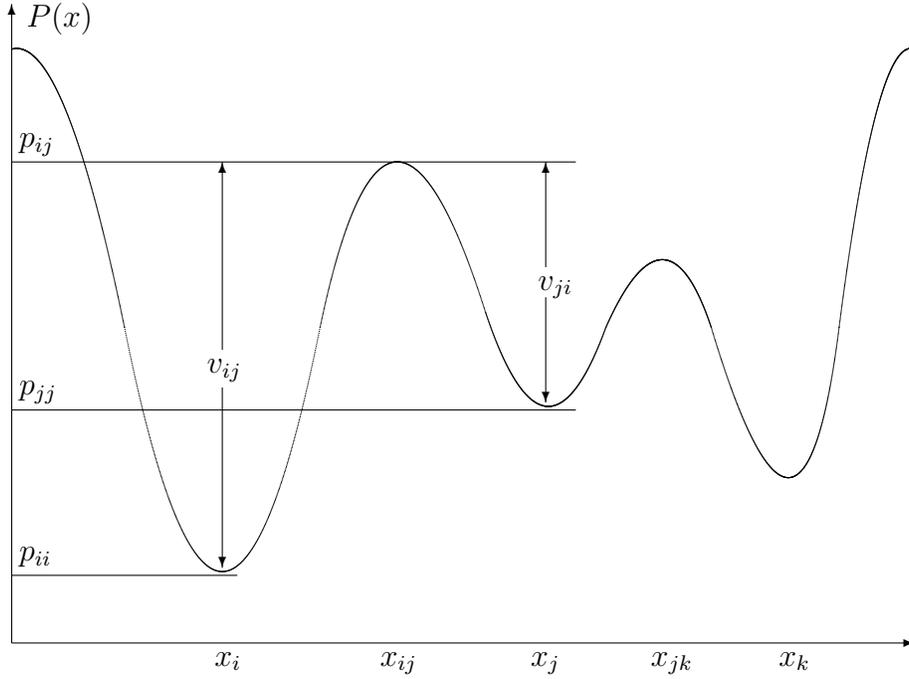


FIG. 2. Construction of an undirected graph P and an directed graph V based on the potential $P(x)$.

6.3. Relationship between the barrier graph V and the potential graph P . Let P be a reflexive (each vertex is incident to a loop) connected weighted undirected graph. We will call such a graph a *potential graph*.

On the same vertex set, each potential graph can be assigned a uniquely weighted directed graph V without loops according to the rule: arc $(i, j) \in AV$, $i \neq j$, if there is an edge $(i, j) \in EP$. The weight of each arc is equal to

$$v_{ij} = p_{ij} - p_{ii} . \quad (30)$$

We will call such an directed graph V *an directed graph of (potential) barriers*. It follows from the definition that the directed graph of barriers is not simply strongly connected, but, moreover, if $(i, j) \in AV$, then $(j, i) \in AV$.

Given a barrier digraph V , the relation (30) does not uniquely determine the potential graph P . If the same number is added to all edge weights (including loops) of the graph P , then the corresponding graph V will be the same. This ambiguity is unimportant for further constructions.

Let us explain the definitions using the example of a physical potential $P(x)$ of sufficient smoothness (see Fig. 2). The minimum points x_i of the function $P(x)$ determine the vertices i of the graph P . The potential values at these points correspond to the weights of the loops, and the values at the saddle points of the x_{ij} between the regions of attraction of the dynamical system $\dot{x} = -\nabla P(x)$, having a common boundary (in one dimension — at points of local maxima), determine the weights of the edges. Exactly

$$p_{ij} = P(x_{ij}) , \quad p_{ii} = P(x_i) .$$

Note also that in the situation of physical potential, the saddle points surrounding vertex i must have a value $P(x)$ higher than the corresponding minimum point. Then all potential barriers are positive ($v_{ij} > 0$, for $(i, j) \in AV$). For the weights of the graph P this means

$$p_{ii} < p_{ij} , \quad (i, j) \in EP , i \neq j .$$

If the last property is not satisfied, then, although the graph P does not correspond to the function $P(x)$ with the corresponding local minima and saddle points (and the arcs of the barrier graph V can have a negative weight), nevertheless this circumstance is not reflected in the properties under study.

Let us note one more property of the weights of the barrier graph V . From (30) we have: $v_{ij} - v_{ji} = p_{jj} - p_{ii}$. Therefore, if an edge is $(i, j) \in EP$, then $v_{ij} < v_{ji}$ if and only if $p_{ii} > p_{jj}$.

If P is an arbitrary connected weighted graph, then it can be turned into a reflexive graph by simply adding the missing loops and assigning arbitrary weights to them. The resulting graph is already a potential graph and is uniquely associated with the barrier directed graph V . However, with respect to the original graph P , such a mapping is

obviously no longer unambiguous and there is arbitrariness associated with the choice of loop weights.

6.4. Potential shift.

Claim 1. *Let two undirected graphs P and P' differ only by a shift of a fixed value d in the weight function: $p_{ij} = p'_{ij} + d$. Let also F and F' be two k -component spanning forests of the graphs P and P' , respectively. And let $\mathbf{EF} = \mathbf{EF}'$. Then the forests F and F' are simultaneously minimal or not.*

Proof. Indeed, an N -vertex undirected forest consisting of k components contains exactly $N - k$ edges. Therefore, for an arbitrary k -component spanning forest of P and the corresponding forest F' (with the same set of edges), which is a subgraph of P' , we have

$$\Upsilon^F - \Upsilon^{F'} = d(N - k) .$$

Thus, when shifting the weight function, k -component forests change their weight by the same fixed value. Therefore, for any two forests F and G of P and the corresponding forests F' and G' of P'

$$\Upsilon^F - \Upsilon^G = \Upsilon^{F'} - \Upsilon^{G'} .$$

□

By virtue of this statement, the shift of the potential does not affect the structure of its minimal forests and trees. To find minimal forests of the barrier graph V , it is sufficient to use any chosen potential graph P , related to V by the relation (30). Therefore, below we fix the potential graph P by an arbitrary representative.

6.5. Removal and giving orientation of trees with replacement of weights. First of all, let us note that if there is an undirected tree T on the set of vertices $\mathbf{S} \subseteq \mathbf{N}$ ($\mathbf{VT} = \mathbf{S}$), then by choosing a vertex $q \in \mathbf{S}$, it can be turned into an entering tree T_q in a unique way, simply by appointing this vertex as the root. The directions of the arcs are determined automatically. The opposite action is even simpler. An directed tree T_q is uniquely transformed into an undirected tree T by simply replacing all arcs with edges. However, in our situation there are two distinguished weighted graphs - the barrier directed graph V and the unoriented potential graph P . Therefore, it is natural to perform the operations of removing and giving orientation with a simultaneous replacement of weights.

Let T_q be a subgraph of the barrier digraph V that is an entering tree with a root at the vertex $q \in \mathbf{S} \subseteq \mathbf{N}$, $\mathbf{VT}_q = \mathbf{S}$ ($T_q \in \mathcal{T}_{\mathbf{S}}^{\bullet q}$). By *removing*

orientation with weight replacement we will mean associating it with an undirected tree T , in which all arcs (r, t) are transformed into edges with a simultaneous replacement of the weight v_{rt} with p_{rt} . Obviously, $T \in \mathcal{T}_S$.

Now let T be a tree of the potential graph P with vertex set $S \subseteq N$: $T \in \mathcal{T}_S$ and $q \in S$. *Giving orientation with weight replacement* is the mapping of the entering tree T_q with the assigned root q to the tree T . In this case, if the arc (i, j) belongs to the obtained tree, its weight is considered to be the number $v_{ij} = p_{ij} - p_{ii}$. Obviously, $T_q \in \mathcal{T}_S^{\bullet q}$. Let us agree to call the trees T and T_q corresponding to each other.

Note that the operation of removing orientation is carried over to forests F of the barrier graph V without changes. If there is an undirected forest F , then to give it orientation it is necessary to specify roots on each connectivity component.

Let us establish a connection between the weights of entering trees on the set $S \subseteq N$ with different roots and the weights of the corresponding undirected trees.

Claim 2. *Let $q \in S \subseteq N$, $T_q \in \mathcal{T}_S^{\bullet q}$, and T be the corresponding undirected tree. Then*

$$\Upsilon^{T_q} = \Upsilon^T - \sum_{t \in S} p_{tt} + p_{qq} . \quad (31)$$

Proof. In the entering tree T_q , arcs from the set S do not outgo and there are no arcs entering to S . Therefore, $\Upsilon^{T_q} = \Upsilon_S^{T_q} = \Upsilon^{T_q|S}$. We have

$$\Upsilon^{T_q} = \sum_{(r,t) \in AT_q} v_{rt} = \sum_{(r,t) \in AT_q} (p_{rt} - p_{tt}) . \quad (32)$$

In the entering tree T_q , a single arc originates from each vertex of the set S except the root q . Therefore, in the last expression, each value of p_{tt} is present exactly once, with the exception of the missing value of p_{qq} . Thus

$$\Upsilon^{T_q} = \sum_{(r,t) \in AT_q} (p_{rt} - p_{tt}) = \sum_{(r,t) \in ET} p_{rt} - \sum_{t \in S \setminus \{q\}} p_{tt} .$$

In the last expression, the first sum corresponds to the weight of the undirected tree T and is equal to Υ^T . Therefore

$$\Upsilon^{T_q} = \Upsilon^T - \sum_{t \in S \setminus \{q\}} p_{tt} . \quad (33)$$

Considering that

$$\sum_{t \in \mathbf{S} \setminus \{q\}} p_{tt} = \sum_{t \in \mathbf{S}} p_{tt} - p_{qq} ,$$

from (33) we obtain (31). □

From the proven statement it immediately follows

Theorem 3. *Let $q \in \mathbf{S} \subseteq \mathbf{N}$, $T_q \in \mathcal{T}_{\mathbf{S}}^{\bullet q}$ and \mathbf{T} be the corresponding undirected tree. Then $T_q \in \tilde{\mathcal{T}}_{\mathbf{S}}^{\bullet q}$ if and only if $\mathbf{T} \in \tilde{\mathcal{T}}_{\mathbf{S}}$.*

Proof. In the expression (33), for a fixed q , the value of $\sum_{t \in \mathbf{S} \setminus \{q\}} p_{tt}$ is the same for all trees in $\mathcal{T}_{\mathbf{S}}^{\bullet q}$. Therefore, the minimum weights of the trees T_q and \mathbf{T} are reached simultaneously. □

A special case of this theorem is considered in [18].

Claim 3. *Let $\mathbf{T} \in \mathcal{T}_{\mathbf{S}}$, and let T_q and $T_{q'}$ be the corresponding entering trees with roots at q and q' , respectively. Then*

$$\Upsilon^{T_q} - p_{qq} = \Upsilon^{T_{q'}} - p_{q'q'} . \quad (34)$$

Proof. Writing the expression (31) for the vertices q and q' and subtracting one from the other, we obtain

$$\Upsilon^{T_q} - \Upsilon^{T_{q'}} = p_{qq} - p_{q'q'} .$$

which coincides with (34). □

6.6. Weights of type λ and type ν .

Claim 4. *Let $\{q, q'\} \subseteq \mathbf{S} \subseteq \mathbf{N}$, $\mathbf{T} \in \tilde{\mathcal{T}}_{\mathbf{S}}$, then*

$$\lambda_{\mathbf{S}}^{\bullet q} - p_{qq} = \lambda_{\mathbf{S}}^{\bullet q'} - p_{q'q'} . \quad (35)$$

Proof. We introduce T_q and $T_{q'}$ — the corresponding trees to the tree \mathbf{T} with roots at q and q' , respectively. Note that according to Theorem 3 $T_q \in \tilde{\mathcal{T}}_{\mathbf{S}}^{\bullet q}$ and $T_{q'} \in \tilde{\mathcal{T}}_{\mathbf{S}}^{\bullet q'}$. Thus, $\lambda_{\mathbf{S}}^{\bullet q} = \Upsilon^{T_q}$ and $\lambda_{\mathbf{S}}^{\bullet q'} = \Upsilon^{T_{q'}}$. Now the statement follows from (34). □

Claim 5. *Let $q \in \mathbf{S} \subseteq \mathbf{N}$, $\mathbf{T} \in \tilde{\mathcal{T}}_{\mathbf{S}}$. Then*

$$\lambda_{\mathbf{S}}^{\bullet q} = \nu_{\mathbf{S}} - \sum_{t \in \mathbf{S}} p_{tt} + p_{qq} . \quad (36)$$

Moreover, if $p_{yy} = \min_{t \in \mathbf{S}} p_{tt}$, then

$$\lambda_{\mathbf{S}}^{\bullet} = \lambda_{\mathbf{S}}^{\bullet y} = \Upsilon^{T_y} , \quad (37)$$

where T_y is rooted at y the entering tree corresponding to the tree \mathbf{T} .

Proof. By condition $\Upsilon^T = \nu_S$. According to Theorem 3, for each of the entering trees T_q corresponding to the tree T , $\lambda_S^{\bullet q} = \Upsilon^{T_q}$ holds. Substituting this value into (31) from Claim 2, we obtain (36).

In (36) the only term that depends on q is the last term p_{qq} . By assumption, it is minimal when $q = y$. Thus

$$\begin{aligned} \lambda_S^{\bullet} &= \min_{q \in S} \lambda_S^{\bullet q} = \min_{q \in S} (\nu_S - \sum_{t \in S} p_{tt} + p_{qq}) = \\ &= \nu_S - \sum_{t \in S} p_{tt} + \min_{q \in S} p_{qq} = \Upsilon^T - \sum_{t \in S} p_{tt} + p_{yy} = \Upsilon^{T_y} . \end{aligned}$$

The last equality takes into account (31). \square

Claim 6. Let $S \subsetneq N$, then

$$\lambda_S^{\circ} = \nu_S - \sum_{t \in S} p_{tt} + \min_{\substack{q \in S, \\ r \notin S}} p_{qr} . \quad (38)$$

Proof. Let's write (16), taking into account (36)

$$\begin{aligned} \lambda_S^{\circ} &= \min_{q \in S} \left(\lambda_S^{\bullet q} + \min_{r \notin S} v_{qr} \right) = \min_{q \in S} \left(\nu_S + p_{qq} - \sum_{t \in S} p_{tt} + \min_{r \notin S} v_{qr} \right) = \\ &= \nu_S - \sum_{t \in S} p_{tt} + \min_{q \in S} \left(p_{qq} + \min_{r \notin S} v_{qr} \right) = \\ &= \nu_S - \sum_{t \in S} p_{tt} + \min_{q \in S} \left(p_{qq} + \min_{r \notin S} (p_{qr} - p_{qq}) \right) = \nu_S - \sum_{t \in S} p_{tt} + \min_{\substack{q \in S, \\ r \notin S}} p_{qr} . \end{aligned}$$

\square

Claim 7. Let $S \subsetneq N$, and $p_{ab} = \min_{\substack{q \in S, \\ r \notin S}} p_{qr}$, then

$$\lambda_S^{\circ} = \lambda_S^{\bullet a} + v_{ab} . \quad (39)$$

Proof. Let's add and subtract the value p_{aa} into (38) and take into account (36):

$$\lambda_S^{\circ} = \nu_S - \sum_{t \in S} p_{tt} + p_{ab} + p_{aa} - p_{aa} = \lambda_S^{\bullet a} + p_{ab} - p_{aa} = \lambda_S^{\bullet a} + v_{ab} .$$

\square

Claim 8. Let $t \in S \subsetneq N$, $p_{ab} = \min_{\substack{q \in S, \\ r \notin S}} p_{qr}$, then

$$\lambda_S^{\circ} = \lambda_S^{\bullet t} - p_{tt} + p_{ab} . \quad (40)$$

Proof. Let's write it down (39) and take it into account (35):

$\lambda_{\mathbf{S}}^{\circ} = \lambda_{\mathbf{S}}^{\bullet a} + v_{ab} = \lambda_{\mathbf{S}}^{\bullet a} + p_{ab} - p_{aa} = \lambda_{\mathbf{S}}^{\bullet a} - p_{aa} + p_{ab} = \lambda_{\mathbf{S}}^{\bullet t} - p_{tt} + p_{ab}$,
which coincides with (40). □

Claim 9. *Equality is fair*

$$\lambda_{\mathbf{S}}^{\circ} - \lambda_{\mathbf{S}}^{\bullet} = \min_{\substack{q \in \mathbf{S}, \\ r \notin \mathbf{S}}} p_{qr} - \min_{t \in \mathbf{S}} p_{tt} . \quad (41)$$

Proof. Let $p_{yy} = \min_{t \in \mathbf{S}} p_{tt}$. According to Claim 5 $\lambda_{\mathbf{S}}^{\bullet} = \lambda_{\mathbf{S}}^{\bullet y}$. The expression (40) takes the form

$$\lambda_{\mathbf{S}}^{\circ} = \lambda_{\mathbf{S}}^{\bullet} - p_{yy} + p_{ab} . \quad (42)$$

Substituting the definitions of p_{yy} and p_{ab} into the quantities, we obtain (41). □

The increment of weights (41) has a natural physical interpretation. It is equal to the value of the potential barrier that must be overcome (or, equivalently, the work that must be done) in order to go beyond the deepest point of the set \mathbf{S} .

7 On the algorithmization of construction of minimal forests and trees of the barrier graph

For arbitrary digraphs, there is a unique algorithm for constructing minimum outgoing forests (up to a spanning minimum tree) [12]. It uses as a subalgorithm the very sophisticated Chu-Liu-Edmonds algorithm [3]-[6] for outgoing trees with a given root ¹, which, in turn, is the only algorithm for constructing minimal directed trees.

Let us address the question of how to use the properties of the weights of the barrier graph V (that is, use the corresponding potential graph P) to construct minimal directed forests and trees.

7.1. Construction of minimal trees of the barrier digraph.

Claims 4 and 5 allow us to easily construct minimal trees of an directed graph V on an arbitrary subset $\mathbf{S} \subseteq \mathbf{N}$ (as long as it is connected), including a spanning minimal tree. Indeed, the Chu-Liu-Edmonds algorithm for constructing a spanning minimal tree with a given root has complexity $O(N^2)$ for dense directed graphs. If it is

¹The Chu-Liu-Edmonds algorithm is originally formulated for outgoing trees. It is also called optimal branching. By reversing the arc directions, we obtain entering trees.

required to find the minimum without regard to the root, then in the general situation it must be applied N times. This increases the complexity of the algorithm and it becomes $O(N^3)$. In our situation for the barrier orgraph it is sufficient to apply this algorithm exactly once. According to Claim 5, the root is immediately assigned to the vertex l with the minimal loop: $p_{yy} = \min_{t \in N} p_{tt}$. In addition, in the resulting tree, by simply reassigning the root, minimal trees with all remaining roots are immediately determined. Their weights are determined according to Claim 4 from formula (35). The complexity of the algorithm remains $O(N^2)$.

Moreover, there is no need to use the Chu-Liu-Edmonds algorithm, which is quite non-trivial in implementation. We can move on to an undirected graph of potential P . For it, use the simple algorithms of Prim [1] or Kruskal [2]. Formally, their complexity for dense graphs is $O(N^2)$. Using any of them, we obtain a minimal spanning undirected tree. Giving it an orientation by simply assigning a root, we again obtain minimal directed trees according to Theorem 3. All their weights are immediately found by Claim 5 from formula (36), which does not increase the overall complexity and it remains $O(N^2)$. If it is necessary to determine a minimal tree regardless of the root, then again the root is immediately assigned to the vertex l with the minimal loop. Thus, the construction of a spanning minimum tree for a barrier digraph is implemented more simply and N times more efficiently than for an arbitrary digraph.

7.2. The problem of constructing minimum spanning forests.

With spanning entering forests the situation is significantly different. Knowing the minimal undirected forest F does not lead to obtaining the corresponding directed minimal forest. It is not difficult to determine the smallest value of p_{yy} on each connected component, and then orient the forest. The resulting forest, unfortunately, is not minimal. The problem is that the connectivity components of undirected and directed minimal forests can be completely different. It is even possible that the vertex set of any tree of any forest from \tilde{F}^k does not coincide with any vertex set of any tree of any directed forest from $\tilde{\mathcal{F}}^k$.

7.3. Example. Let us consider an example (see Fig. 3) demonstrating a significant difference between the minimal entering forests of the barrier graph V and the minimal forests of the corresponding potential graph P . Here the weights of the arcs V are equal to: $v_{ab} = 1$, $v_{ba} = 4$, $v_{bc} = 3$, $v_{cb} = 2$. The potential graph P has edge (and loop) weights: $p_{ab} = 5$, $p_{bc} = 4$, $p_{aa} = 4$, $p_{bb} = 1$, $p_{cc} = 2$. As

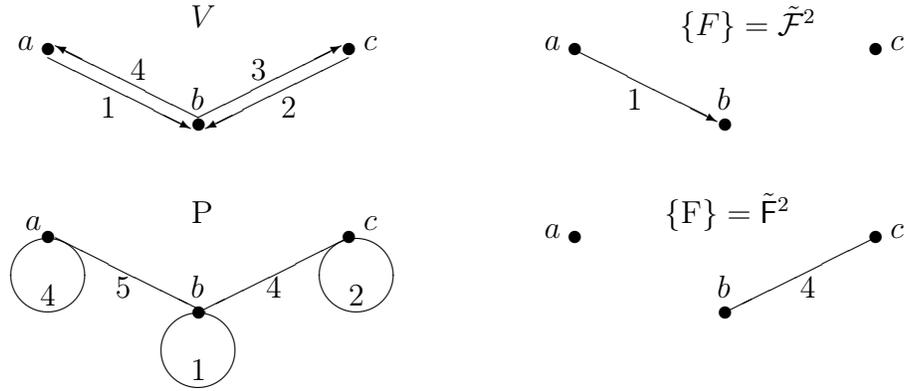


FIG. 3. Above is the barrier graph V and its spanning minimum entering forest F , consisting of two trees; below is the corresponding potential graph P and its spanning minimum forest F , consisting of two trees.

is easy to see, (30) holds for V and P . The minimum spanning forest F of a digraph V consisting of two trees is unique. It consists of an empty tree rooted at c and a tree containing a single arc (a, b) with weight 1. The minimum spanning forest F consisting of two trees is also unique. But it contains an empty tree consisting of one vertex a and a tree consisting of one edge (b, c) of weight 4. Thus, the vertex set of any tree in the entering forest F does not coincide with any tree vertex set in the forest F .

This example demonstrates that knowing the minimal forests of the potential graph P does not help in finding the minimal entering forests of the corresponding barrier digraph V . This is quite natural, since the forests of the graph P do not contain loops. Moreover, each loop (q, q) of the graph P participates in the weight of all arcs outgoing from vertex q in the directed graph V . Because of this, other sets of tree vertices are formed. In the case of spanning trees (or trees on a given subset of a set of vertices), this problem does not exist, since the set of tree vertices is known in advance. And on the selected subset, if we set the root, in the weight of any entering tree, according to (31), the same sum of loops of the graph P will appear. Therefore, the type of the minimal entering tree ceases to depend on loops.

7.4. Guiding Considerations. Let there be some minimal spanning entering forest $F \in \tilde{\mathcal{F}}^k$. Then the weights of its trees T_l^F , $l \in K_F$ are known. According to Property 3, they are equal to $\gamma^{T_l^F} = \lambda_{VT_l^F}^\bullet = \lambda_{VT_l^F}^\bullet$. By the way, for any set $S \subseteq N$, the value

$\lambda_{\mathfrak{S}}^\bullet$, according to Claim 5, is easily determined ((36), (37)) from the potential graph P. For it, the corresponding minimum $\nu_{\mathfrak{S}}$ is effectively found using Prim's or Kruskal's algorithm on any set of vertices \mathfrak{S} .

According to Theorem 2, to construct the minimal descendant of the entering forest F , it is necessary to know the values $\lambda_{\mathfrak{V}T_l^F}^\circ$. Then it is necessary to determine the increments $\lambda_{\mathfrak{V}T_l^F}^\circ - \lambda_{\mathfrak{V}T_l^F}^\bullet$ (see (41)) and find the smallest among them (let, for definiteness, the smallest increment at $l = y$). After this, the arcs of the incoming tree T_y^F are replaced by the arcs of the tree T on which the minimum of $\lambda_{\mathfrak{V}T_y^F}^\circ$ is achieved. If in the tree T the arc that comes from $\mathfrak{V}T_y^F$ (and it is the only one) is the arc (a, b) , then in fact in the tree T_y^F the root is reassigned to the vertex a and the arc (a, b) is added. As a result, we obtain the minimal entering forest $G \in \tilde{\mathcal{F}}^{k-1}$, which is a descendant of the entering forest F .

According to Theorem 2 (21) and Claim 9 (41), the problem is to determine the minimum of the increment

$$\min_{l \in \mathfrak{K}_F} (\lambda_{\mathfrak{V}T_l^F}^\circ - \lambda_{\mathfrak{V}T_l^F}^\bullet) = \min_{l \in \mathfrak{K}_F} \left(\min_{\substack{q \in \mathfrak{V}T_l^F, \\ r \notin \mathfrak{V}T_l^F}} p_{qr} - \min_{t \in \mathfrak{V}T_l^F} p_{tt} \right). \quad (43)$$

With the definition of values $\min_{t \in \mathfrak{V}T_l^F} p_{tt}$ there is no problem. They are equal to the weights of the loops p_{ll} of the potential graph P at the roots of the corresponding trees of the entering forest F . So they are known in advance. The expression (43) takes the form

$$\min_{l \in \mathfrak{K}_F} (\lambda_{\mathfrak{V}T_l^F}^\circ - \lambda_{\mathfrak{V}T_l^F}^\bullet) = \min_{l \in \mathfrak{K}_F} \left(\min_{\substack{q \in \mathfrak{V}T_l^F, \\ r \notin \mathfrak{V}T_l^F}} p_{qr} - p_{ll} \right). \quad (44)$$

Thus, the problem is reduced to determining the minimal barriers between the set of vertices of a particular tree and the remaining vertices.

7.5. Barriers between sets of tree vertices. We transform the expression (44) to introduce minimal barriers between different trees:

$$\min_{l \in \mathfrak{K}_F} \left(\min_{i \in \mathfrak{K}_F \setminus \{l\}} \min_{\substack{q \in \mathfrak{V}T_i^F, \\ r \in \mathfrak{V}T_l^F}} p_{qr} - p_{ll} \right) = \min_{\{l, i\} \subset \mathfrak{K}_F} \left(\min_{\substack{q \in \mathfrak{V}T_i^F, \\ r \in \mathfrak{V}T_l^F}} p_{qr} - p_{ll} \right). \quad (45)$$

The value in brackets now corresponds to the potential barrier that must be overcome to get from the deepest point of the tree T_l^F , namely

the vertex l , to the set of vertices of the tree rooted at i . And the question now is how to efficiently count transitions between trees.

Let $F \in \tilde{\mathcal{F}}^k$. We introduce an enlarged directed graph V^k of barriers between the sets of vertices of the trees of the forest F . The set of vertices is the roots of the trees: $\mathcal{V}V^k = \mathcal{K}_F$. We define the weights v_{li}^k of the arcs as

$$v_{li}^k = \min_{\substack{q \in \mathcal{V}T_l^F, \\ r \in \mathcal{V}T_i^F}} p_{qr} - p_{li}, \quad l \neq i. \quad (46)$$

If in the directed graph V there are no arcs between the sets of vertices of the trees T_l^F and T_i^F , then we assume that in the directed graph V^k there are no arcs (l, i) and (i, l) .

Similarly, the enlarged graph of the potential P^k is defined, for which $\mathcal{V}P^k = \mathcal{V}V^k$. The weights of the loops remain the same $p_{li}^k = p_{li}$. The weights of the edges are equal to

$$p_{li}^k = \min_{\substack{q \in \mathcal{V}T_l^F, \\ r \in \mathcal{V}T_i^F}} p_{qr}, \quad l \neq i. \quad (47)$$

If the graph P does not contain edges between the sets of vertices of the trees T_l^F and T_i^F , then we assume that the enlarged graph P^k does not contain the edge (l, i) .

By construction, the introduced directed graph V^k is a barrier directed graph, and P^k is the corresponding potential graph.

Let the arc of the minimum weight of the enlarged barrier graph V^k be the arc (y, x) , and let (a, b) be that edge of the original potential graph for which

$$p_{ab} = \min_{\substack{q \in \mathcal{V}T_x^F, \\ r \in \mathcal{V}T_y^F}} p_{qr} = p_{yx}^k, \quad a \in \mathcal{V}T_y^F, \quad b \in \mathcal{V}T_x^F. \quad (48)$$

Thus,

$$\min_{(l,i) \in \mathcal{A}V^k} v_{li}^k = v_{yx}^k = p_{yx}^k - p_{yy}^k = p_{ab} - p_{yy}. \quad (49)$$

In this case, $a \in \mathcal{V}T_y^F$, $b \in \mathcal{V}T_x^F$, and the value (49) by construction coincides with both (44) and (41), where $\mathcal{S} = \mathcal{V}T_y^F$, and therefore with (23). Thus, adding an arc (a, b) to the forest F and reorienting the tree T_y^F by assigning the root to the vertex a , we obtain an encroaching forest $G \in \tilde{\mathcal{F}}^{k-1} \cap \mathcal{R}^F$. For it, according to Proposition 2 (27) and Propositions 8 (40) and 9 (42), the weight of the increased tree changes (the tree with the root at x absorbed the tree with the root at y):

$$\lambda_{\mathbf{V}T_x^G}^\bullet = \lambda_{\mathbf{V}T_x^F}^\bullet + \lambda_{\mathbf{V}T_y^F}^\circ = \lambda_{\mathbf{V}T_x^F}^\bullet + \lambda_{\mathbf{V}T_y^F}^\bullet + p_{ab} - p_{yy} .$$

Accordingly, the weight of the new minimum forest is equal to

$$\mathcal{Y}^G = \mathcal{Y}^F + p_{ab} - p_{yy} . \quad (50)$$

7.6. Recalculation of barriers and potential. The resulting digraph G is minimal and consists of $k - 1$ trees. Let us create an enlarged digraph of barriers V^{k-1} for it. In the digraph of barriers V^{k-1} (and the potential graph \mathbf{P}^{k-1}), compared to V^k (and \mathbf{P}^k), the vertex y has disappeared from the set of vertices. The weights of arcs (edges) not incident to vertex x have remained the same.

$$v_{li}^{k-1} = v_{li}^k, p_{li}^{k-1} = p_{li}^k, \quad \{l, i\} \subset K_G \setminus \{x\} . \quad (51)$$

The loop weights of the potential graph \mathbf{P}^{k-1} for $l \neq x$ are obviously the same. The weight of the loop (x, x) has not changed either, since this vertex remains in the graph.

$$p_{tt}^{k-1} = p_{tt}^k, \quad t \in \mathbf{V}\mathbf{P}^{k-1} = \mathbf{V}\mathbf{P}^k \setminus \{y\} . \quad (52)$$

For the weights of the edges of the barrier graph \mathbf{P}^{k-1} incident to the vertex x , we have

$$p_{lx}^{k-1} = \min_{\substack{q \in \mathbf{V}T_l^G, \\ r \in \mathbf{V}G_x^F}} p_{qr} = \min_{\substack{q \in \mathbf{V}T_l^G, \\ r \in \mathbf{V}T_x^F \cup \mathbf{V}T_y^F}} p_{qr} = \min_{t \in \{x, y\}} \min_{\substack{q \in \mathbf{V}T_l^F, \\ r \in \mathbf{V}F_t^F}} p_{qr}, \quad l \in \mathbf{V}\mathbf{P}^{k-1} \setminus \{x\},$$

where do we get that

$$p_{lx}^{k-1} = \min(p_{lx}^k, p_{ly}^k), \quad l \neq x . \quad (53)$$

After this, the weights of the corresponding arcs of the barrier digraph V^k are immediately determined

$$v_{lx}^{k-1} = p_{lx}^{k-1} - p_{ll}^{k-1}, \quad v_{xl}^{k-1} = p_{lx}^{k-1} - p_{xx}^{k-1}, \quad l \in \mathbf{V}\mathbf{P}^{k-1} \setminus \{x\} . \quad (54)$$

7.7. Algorithm for constructing minimum entering forests and its complexity. Zero step (start of the algorithm).

At this step, there are weight matrices $\mathbf{V}^N = \mathbf{V}$ and $\mathbf{P}^N = \mathbf{P}$ of the barrier graph V and the potential graph \mathbf{P} , respectively. In the absence of an arc (edge), the weight is interpreted as ∞ . The initial minimal forest is a spanning forest $F \in \tilde{\mathcal{F}}^N$, which is empty. It has no arcs. Its weight is $\mathcal{Y}^F = 0$.

Lists Ω_t of edges (the orientation within the list is unimportant, so we use the term edge rather than arc) of the trees of the minimal forest are created. These lists are indexed by the roots. Knowing the edges and roots, it is easy to determine what an entering forest looks like. At the initial stage, there are N of these lists, and they are all empty.

Next, the procedure is recurrent.

l -th step ($l = 1, 2, \dots, N - 1$).

The currently constructed minimal entering forest F contains $k = N + 1 - l$ roots. It is defined by the current lists Ω_t of edges and the roots $t \in K_F$. The current matrix \mathbf{V}^k of barriers and the matrix \mathbf{P}^k of potentials are known.

a) The minimum element in the barrier matrix \mathbf{V}^k is determined. The roots corresponding to this element are given the labels y and x :

$$v_{yx} = \min_{(l,i) \in \mathbf{V}^k} v_{li}^k .$$

b) The vertices of the edge of the original potential graph corresponding to the element p_{yx}^k receive the labels a and b :

$$p_{ab} = p_{yx}^k , \quad a \in \mathbf{V}T_y^F , \quad b \in \mathbf{V}T_x^F .$$

c) The tree T_y^F is reoriented by assigning a root at a . An arc (a, b) is added to it (with its weight v_{ab}). This turns the forest F into the new minimum entering forest under study belonging to $\tilde{\mathcal{F}}^{k-1}$. The tree rooted at x absorbs the tree rooted at y . We note the transformations necessary for this. According to (50), the new weight of the minimum forest is

$$\mathcal{Y}^F := \mathcal{Y}^F + p_{ab} - p_{yy} .$$

The edge lists Ω_x and Ω_y are combined into a single list Ω_x and the edge (a, b) is added to it. This completely defines the new tree T_x^F and the new minimal entering forest F itself.

d) Matrices \mathbf{V}^{k-1} and \mathbf{P}^{k-1} are constructed. For the edge weight matrix, $k-2$ comparisons (53) are required to be performed to select the minimum ones. They are used to adjust the current edge matrix of the potential. At the same time, the edges of the original potential graph that yield these minima (which are used in point b) are remembered. The modified $2(k-2)$ arcs (54) are calculated and the matrix of arc weights is adjusted. Now the index corresponding to the vertex y is removed from the matrices. It is no longer the root of the tree, and the dimension of the matrices is reduced by 1 and becomes equal to $k-1$. The matrices \mathbf{V}^{k-1} and \mathbf{P}^{k-1} are obtained.

e) We set $l := l + 1$. If $l = N$, then end. Otherwise, go to point a).

Property 4, common to directed graphs, and Theorem 2, taking into account the equality of the formula (23) in it to the value (49), which is fully used in the algorithm, guarantee that the proposed procedure can construct any minimal forest from $\tilde{\mathcal{F}}^k$, $k = 1, 2, \dots, N$.

7.8. Complexity of the algorithm. When estimating the complexity, we assume that the original graph is dense $|AV| \sim N^2$.

a) Selecting the minimum element at each step with the current number of roots k (the matrix size is $k \times k$) takes $O(k^2)$ operations. Summing up over all steps, we get the complexity of this most expensive item $O(N^3)$.

d) Comparison and correction of arc weights requires $O(k)$ operations. Summing up over all steps, we obtain the complexity of this point $O(N^2)$.

The total complexity is therefore equal to $O(N^3)$.

8 More about the properties of minimal forests of the barrier digraph

According to Theorem 3, the set of minimal incoming forests created as a result of the algorithm has the following property by construction. Let $F \in \tilde{\mathcal{F}}^n$ and $F' \in \tilde{\mathcal{F}}^m$, $n > m$, and they are obtained by the above procedure. When the orientation is removed, they turn into some spanning forests F and F' of the potential graph P , respectively (they are, generally speaking, no longer minimal, as the example considered shows, although each individual tree is minimal on its set of vertices). Then $EF \subset EF'$. An arbitrary directed graph V does not have such a property for its minimal forests when removing the orientation (regardless of the weights themselves), regardless of the algorithm by which they were obtained.

If we abstract from the algorithm and talk about the entering forests themselves, then the described property can be formulated as follows. Let $n > m$, then for any $F \in \tilde{\mathcal{F}}^n$ there is an $F' \in \tilde{\mathcal{F}}^m$ (and vice versa, for any $F' \in \tilde{\mathcal{F}}^m$ there is an $F \in \tilde{\mathcal{F}}^n$) such that when removing the orientation for the resulting undirected forests P and P' , respectively, $EF \subset EF'$.

9 Road network for cities of varying importance

To illustrate the usefulness of Prim and Kruskal's algorithms, an example is usually given with a network of roads that need to connect N cities. The cost weight of a direct connection between city i and

city j is known and equal to p_{ij} . An undirected graph P with such edge weights is created and its minimum spanning tree is found. We will use the same example to demonstrate the significant advantages of minimum spanning directed forests.

The real situation is far from ideal and it may well be that funds can only be allocated to build l roads. And it remains to be seen which cities will be connected in order to spend the minimum amount of funds. For now, this does not complicate the task much. The point is that the Kruskal algorithm procedure at each step builds a spanning forest of minimum weight, which completely solves the issue of building l roads out of $N - 1$ required. And when new funds appear, it is known where to continue construction.

But the reality is still more complicated. The fact is that the cities themselves are obviously unequal in importance. Therefore, when assessing the effectiveness of construction costs, it is necessary to subtract some values of cities that are connected by roads from the total direct costs p_{ij} . If you immediately select the k most important cities and connect only them, then this will obviously not be optimal. And this is not only because they may be too far from each other. If additional funds appear, then the connection with the cities further in importance will become less and less optimal. Let's say the city was located near an already built road between more important objects, and now it is necessary to build a road to it from one of them, instead of immediately passing through it during the previous construction. Moreover, it is essential that the sequence in which the construction is carried out is optimal. After all, funds may suddenly run out and they will not be enough for the planned l roads, but we would like the costs for those already built to be effective.

The next natural idea is to subtract the value of the linked cities from the direct costs at each step in the utility assessment. This looks appealing, but the devil is in the details. Indeed, if you connect two cities, each of which is already connected to someone, then there is no need to subtract the value of these cities at all (it is already taken into account). But if you connect cities that have not yet been affected by the connection, then you need to subtract the value of both of them. And finally, if one of the cities is already connected to someone, and the second is not connected to anyone, then only the value of the latter should be subtracted. This indicates that with this approach, there is no single regular procedure by which the efficiency of costs at each step would be consistently studied.

And that's where minimal entering forests and algorithms for their construction come to the rescue. Indeed, the values of cities are declared

by the weights of the loops p_{ii} . A graph of potential P and a digraph of barriers V with weights of arcs (30) is created (the weight of each arc $v_{ij} = p_{ij} - p_{ii}$ it makes sense for the effective cost of connecting the cities of i and j , taking into account the value of the city of i). After which the algorithm above sequentially determines the minimum spanning entering forests of the barrier graph V . With this approach, at each step, when calculating the effective utility, the value of exactly one city out of two connected by the road will be taken into account. Effective costs are minimal. Note that according to the specified procedure, if all N cities are connected, then by virtue of Theorem 3, direct costs are also minimal.

10 Matrix minors and forests

Classically [20] by directed trees it is customary to understand outgoing trees (or those growing from the root, as well as rooted). The same applies to algorithms. For example, all the main variants of the algorithm for the minimum spanning tree [3]-[6] build exactly an outgoing tree. However, when describing the minors of matrices, the first to emerge were exactly the opposite — the entering forests [15]. Then, in the spectra of directed graphs, namely, in the study of the Laplace matrix (a matrix with a zero sum over a row) of the directed graph: $l_{ij} = -v_{ij}$, $j \neq i$; $l_{ii} = -\sum_{j \neq i} l_{ij}$, entering forests arise

[16]. Further, after the analysis of the pseudo-potential, entering forests themselves arise again in the lemmas on Markov chains in [17]. In all these works, entering forests are called W -graphs (W is the set of vertices from which arcs do not outgo). When deriving the formula for all minors [21, 22], the digraph is immediately called a forest, which in strict notation is an entering forest. For the standard Laplace matrix of an digraph, no preliminary actions are required, and in the study of its spectrum and eigenspaces, entering forests arise automatically [7]-[9].

Using out-forests in matrix analysis requires a bit more work. First, the arcs of the directed graph are reversed, and the Laplace matrix (with zero sum per row) is created for the resulting directed graph. In its analysis, but in the notation of the weights of the original directed graph (without reorienting the arcs), out-forests arise. This matrix itself, with this approach, is called the Kirchhoff matrix ($l_{ij} = -v_{ji}$, $j \neq i$; $l_{ii} = -\sum_{j \neq i} l_{ij}$) [23]. This method of transition somewhat confuses the situation.

When studying the properties of an directed graph through the matrices corresponding to it, it would be worthwhile to use a unified approach to creating these matrices, without special tricks. It seems more natural to follow the historical approach and names [19]. Namely, a weighted adjacency matrix \mathbf{V} of an directed graph V without loops with elements v_{ij} is created. Further, by analogy with undirected graphs, it is proposed to create two diagonal matrices of half-degrees (valences) of entry and exit \mathbf{D}^{in} and \mathbf{D}^{out} with diagonal elements $d_{ii}^{out} = \sum_{j \neq i} v_{ij}$, $d_{ii}^{in} = \sum_{j \neq i} v_{ji}$ (in undirected graphs this is one matrix of degrees). After this (also by analogy with undirected graphs), the matrices of incoming and outgoing conductivities are determined: $\mathbf{C}^{in,out} = \mathbf{D}^{in,out} - \mathbf{V}$ (in the case of undirected graphs, this is the same conductivity matrix). With this approach, the Laplace matrix turns out to be the matrix of outgoing conductivities \mathbf{C}^{out} (it is precisely the one whose sum of elements by rows is equal to zero) and its analysis is described by incoming forests. But the matrix of incoming conductivities \mathbf{C}^{in} has a zero sum by columns, and outgoing forests arise in its analysis. This approach is completely symmetrical with respect to incoming and outgoing forests. Moreover, the weights of the arcs of the directed graph can be complex and this will not affect the formulas in any way. The matrix of incoming conductivities itself can naturally be considered a Kirchhoff matrix. There is no need to forcibly transpose it so that it also turns out to be a Laplace-type matrix.

Note also that the forest structure of an arbitrary directed graph may differ significantly for its incoming and outgoing forests. So the selection of incoming forests still suffers from some one-sidedness.

For purely directed graph problems (not matrix type), replacing arcs with inverse ones does not change anything significant in principle. Therefore, the use of incoming or outgoing forest variants is more related to the specifics of a particular problem statement: is it more important to you from where or to?

Of course, all the results of this paper and the papers [10]-[12] on which it is based can be reformulated for out-forests. This will entail some modification of definitions. Let's say that the weight of a subgraph of an directed graph V on a subset \mathbf{S} of a vertex set (1) should be defined not as the sum of the weights of the arcs outgoing from the vertices of the set \mathbf{S} , but as the sum of the weights of the arcs that enter \mathbf{S} . In general, the transition to outgoing forests is not particularly difficult, and the description in them is completely equivalent to the description in incoming forests. But in addition to the fact that the characteristics of the Laplace matrices are described by incoming forests, such forests

are more natural in the energy interpretation, so they are used in this work.

References

- [1] R. C. Prim, *Shortest connection networks and some generalizations*, Bell System Technical Journal, **36** (1957), 1389–1401.
- [2] Joseph. B. Kruskal, *On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem*, Proc. AMS., **7**:1 (1956), 48–50.
- [3] Y.J. Chu and T.H. Liu, *On the shortest arborescence of a directed graph*, Sci. Sinica, **14** (1965), 1396–1400.
- [4] J. Edmonds, *Optimum branchings*, J. Res. Nat. Bur. Standards, **71**:B (1967), 233–240.
- [5] R. E. Tarjan, *Finding optimum branchings*, Networks, **7**:1 (1977).
- [6] H.N.Gabov, Z.Galil, T.Spencer, R.E.Tarjan, *Efficient Algorithms for Finding Minimum Spanning Trees in Undirected and Directed Graphs*, Combinatorica, **6**:2 (1986), 109–122.
- [7] V. A. Buslov, *On characteristic polynomial coefficients of the Laplace matrix of a weighted digraph and all minors theorem*, J. Math. Sci. (N. Y.), **212**:2 (2016), 643–653.
- [8] V. A. Buslov *On characteristic polynomial and eigenvectors in terms of tree-like structure of the graph*, J. Math. Sci. (N. Y.), **232**:1 (2018), 6–20.
- [9] V. A. Buslov *On the relationship between multiplicities of the matrix spectrum and signs of components of its eigenvector in a tree-like structure*, J. Math. Sci. (N. Y.), **236**:5 (2019), 477–489.
- [10] V.A. Buslov, *Structure of Minimum-Weight Directed Forests: Related Forests and Convexity Inequalities*, J. Math. Sci. (N. Y.), **247**:1 (2020), 383–393.
- [11] V. A. Buslov, *The Structure of Directed Forests of Minimal Weight: Algebra of Subsets*, J. Math. Sci. (N. Y.), 2021, **255**, pp. 1-16.
- [12] V.A. Buslov, *Algorithm for sequential construction of spanning minimal directed forests*, Journal of Mathematical Sciences (N. Y.), **275**:2 (September 2023), 117–129.
- [13] V. A. Buslov, K. A. Makarov, *Hierarchy of time scales in the case of weak diffusion*, Theoret. and Math. Phys., **76**:2 (1988), 818–826.
- [14] V. A. Buslov, K. A. Makarov, *Lifetimes and lower eigenvalues of an operator of small diffusion*, Math. Notes, **51**:1 (1992), 14–21.
- [15] M.Fiedler, J.Sedláček, *O w-basích orientirovaných grafu*, Časopis Pěst. Mat., **83** (1958), 214–225.
- [16] A. D. Ventsel *On the asymptotic of eigenvalues of matrices with elements of order $\exp\{-V_{ij}/2\varepsilon^2\}$* , Reports of the USSR Academy of Sciences, **202**:2 (1972), 263–266.
- [17] A. D. Ventsel and M. I. Freidlin, *Fluctuations in Dynamic Systems Under Small Random Disturbance*, Moscow, (1979), 429 p.
- [18] V. A. Buslov, M. C. Bogdanov, V. A. Hudobahshov, *On the minimum spanning tree for digraphs with potential weights*, Vestnik of St. Petersburg State University, **10**:3 (2008).

- [19] D. Cvetkovič , M.Doob, H.Sachs, *Spectra of Graphs: Theory and Application*, VEB Deutscher Verlag der Wissenschaften, Berlin, (1980), 368 p.
- [20] W. T. Tutte, *Graph Theory*, Cambridge University Press, 2001, 333 p.
- [21] Chaiken S., *A combinatorial proof of the all minors matrix tree theorem*, SIAM Journal on Algebraic and Discrete Methods, **3:3** (1982), 319–329.
- [22] Moon J. W., *Some determinant expansions and the matrix-tree theorem*, Discrete Mathematics, **124** (1994), 163–171.
- [23] Chebotarev P. Y., Agaev R. P., *Matrix forest theorem and Laplacian digraph matrices*, LAP LAMBERT Academic Publishing GmbH Co.Kg, (2011), 266 p.

VASILY ANATOLIEVICH BUSLOV
ST. PETERSBURG STATE UNIVERSITY,
OLD PETERHOF, ST. ULYANOVSKAYA, 3,
198504, ST. PETERSBURG, RUSSIA
Email address: vbuslov@spbu.ru , abvabv@bk.ru