COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA **14** (1) (1994) 1–21

MONOTONE PATH SYSTEMS IN SIMPLE REGIONS

K. CAMERON and H. SACHS

Received March 1, 1988
Revised September 25, 1992

A monotone path system (MPS) is a finite set of pairwise disjoint paths (polygonal areas) in the xy-plane such that every horizontal line intersects each of the paths in at most one point. A MPS naturally determines a "pairing" of its top points with its bottom points. We consider a simple polygon Δ in the xy-plane wich bounds the simple polygonal (closed) region $\mathbf{D}$. Let $\boldsymbol{T}$ and $\boldsymbol{B}$ be two finite, disjoint, equicardinal sets of points of $\mathbf{D}$. We give a good characterization for the existence of a MPS in $\mathbf{D}$ which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$, and a good algorithm for finding such a MPS, and we solve the problem of finding all MPSs in $\mathbf{D}$ which pair $\boldsymbol{T}$ with $\boldsymbol{B}$. We also give sufficient conditions for any such pairing to be the same.

1. Introduction

All of this paper takes place in the xy-plane. We will freely use the words up, down, left, right, horizontal, highest, etc.

We consider a simple polygon Δ in the xy-plane. Δ and its interior is called a polygonal region, which we will denote by $\mathbf{D}$.

A *monotone path* π is a polygonal arc of positive length such that every horizontal line intersects π in at most one point. A monotone path *in* $\mathbf{D}$ is a monotone path whose interior is contained in the interior of $\mathbf{D}$. A *monotone path system* (MPS) Π (in $\mathbf{D}$) is a finite set of pairwise disjoint monotone paths (in $\mathbf{D}$). The sets of top points and bottom points of paths in Π are denoted by $\boldsymbol{T}(\Pi)$ and $\boldsymbol{B}(\Pi)$, respectively.
A MPS Π naturally determine a pairing of $\boldsymbol{T}(\Pi)$ with $\boldsymbol{B}(\Pi)$.

For the sake of conciseness, we restrict our investigations to MPSs *in* $\mathbf{D}$, but most of the results can easily be extended to systems of monotone paths of $\mathbf{D}$ which are allowed to contain points of Δ in their interior.

Theorem 1.1. *Given point sets, $\boldsymbol{T}$ and $\boldsymbol{B}$, of polygon Δ, which contain no vertices of Δ. If there is a monotone path system in $\mathbf{D}$ which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$, then any such pairing is the same.*

Theorem 1.1 need not be true if $\boldsymbol{T}$ or $\boldsymbol{B}$ contain points in the interior of $\mathbf{D}$ (see Figure 1.0) or vertices of Δ (see Figure 1.1).

The first author's research is supported by the Natural Sciences and Engineering Research Council of Canada

AMS subject classification codes (1980): 52 A 37; 05 C 38; 05 C 70

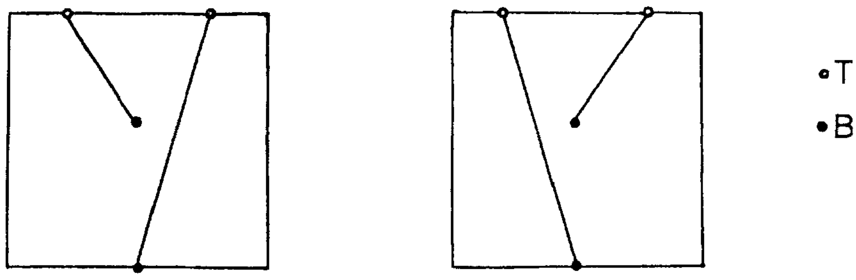

Fig. 1.0. Two different pairings of $\boldsymbol{T}$ with $\boldsymbol{B}$.

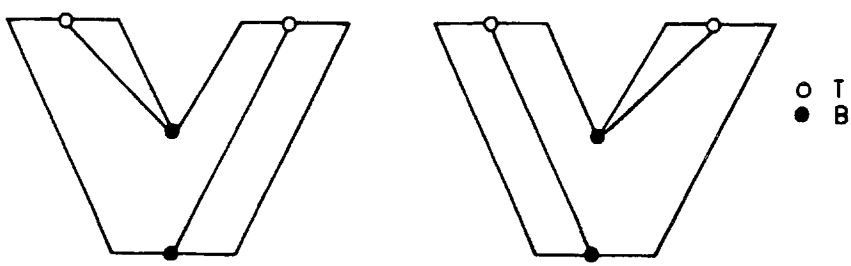

Fig. 1.1. Two different pairings of $\boldsymbol{T}$ with $\boldsymbol{B}$.

Given two sets, $\boldsymbol{T}$ and $\boldsymbol{B}$, of points of polygonal region $\mathbf{D}$, we now consider the problem of finding a MPS Π in $\mathbf{D}$ which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$. It is clear that no such MPS exists unless each of the following conditions holds, so we assume that they do hold through the rest of this paper.

(1) Assumptions:

(1.1) $\boldsymbol{T}$ and $\boldsymbol{B}$ are two finite sets of points of $\mathbf{D}$.

(1.2) $|\boldsymbol{T}|=|\boldsymbol{B}|$.

(1.3) $\boldsymbol{T}\cap\boldsymbol{B}=\emptyset$.

(1.4) No point of $\boldsymbol{T}$ is locally lowest in $\mathbf{D}$ and no point of $\boldsymbol{B}$ is locally highest in $\mathbf{D}$.

Region $\mathbf{D}$ with the points of $\boldsymbol{T}\cup\boldsymbol{B}$ inserted will be denoted by $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$.

A *horizontal chord* $\boldsymbol{C}$ of $\mathbf{D}$ is a horizontal line segment of positive length whose interior is interior to $\mathbf{D}$ and whose end points are on Δ. Horizontal chord $\boldsymbol{C}$ partitions $\mathbf{D}$ into three disjoint parts: $\boldsymbol{C}$; the upper component, $\mathbf{U}(\boldsymbol{C})$, of $\mathbf{D}-\boldsymbol{C}$, which is the component of $\mathbf{D}-\boldsymbol{C}$ for which $\boldsymbol{C}$ is a set of locally lowest points of $\mathbf{U}(\boldsymbol{C})\cup\boldsymbol{C}$; and the lower component, $\mathbf{L}(\boldsymbol{C})$. Define

$$\begin{aligned} d(\boldsymbol{C}) := &\ |\mathbf{U}(\boldsymbol{C})\cap\boldsymbol{T}| - |(\mathbf{U}(\boldsymbol{C})\cup\boldsymbol{C})\cap\boldsymbol{B}| \\ = &\ |\mathbf{L}(\boldsymbol{C})\cap\boldsymbol{B}| - |(\mathbf{L}(\boldsymbol{C})\cup\boldsymbol{C})\cap\boldsymbol{T}| \end{aligned}$$

using (1.2)

Theorem 1.2. *Where $\boldsymbol{T}$ and $\boldsymbol{B}$ contain no vertices of the boundary Δ of $\boldsymbol{B}$, there is a monotone path system Π in $\mathbf{D}$ which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$ if and only if*

(2) $$d(\boldsymbol{C}) \geq 0, \quad \textit{for every horizontal chord } \boldsymbol{C} \textit{ of } \mathbf{D}.$$

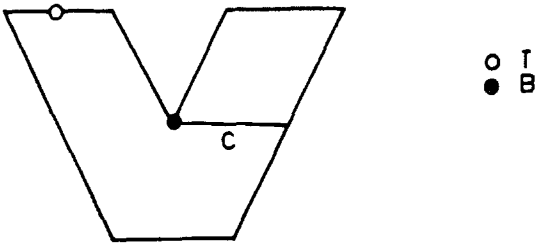

Fig. 1.2. $d(C) = -1$, but there is a pairing of $\boldsymbol{T}$ with $\boldsymbol{B}$.

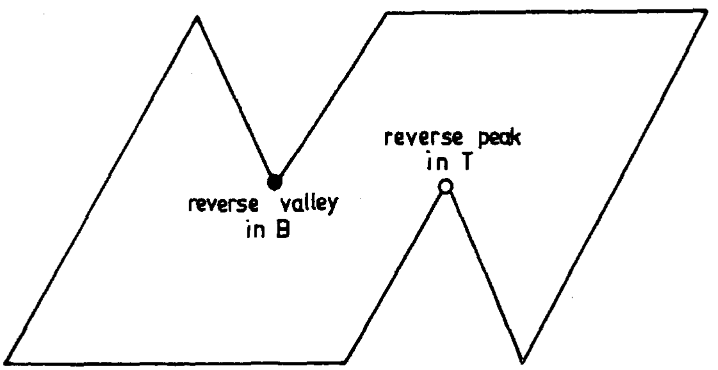

Fig. 1.3. Problem points.

Figure 1.2 shows that if vertices of Δ are allowed to be points of $\boldsymbol{T}$ or $\boldsymbol{B}$, then Theorem 1.2 (the "only if" part) need not hold.

Actually, we do not need to exclude all vertices of Δ from being points of $\boldsymbol{T}$ and $\boldsymbol{B}$ in order for Theorems 1.1 and 1.2 to hold: we only have to exclude "problem points" which we now describe.

A *peak (valley)* is a unique locally highest (lowest) point of $\mathbf{D}$. A *reverse peak (reverse valley) p* is a vertex of Δ, which is unique locally highest (lowest) point of Δ, but not a locally highest (lowest) point of $\mathbf{D}$. A *problem* point is a reverse peak in $\boldsymbol{T}$ or a reverse valley in $\boldsymbol{B}$. See Figure 1.3.

Theorem 1.1'. *Given point sets, $\boldsymbol{T}$ and $\boldsymbol{B}$, of polygon Δ, which contain no problem points. If there is a monotone path system in* $\mathbf{D}$ *which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$, then any such pairing is the same.*

A *distinguished horizontal chord* (DHC) of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ is a horizontal chord of $\mathbf{D}$ containing at least one point which is either in $\boldsymbol{T}\cup\boldsymbol{B}$ or is a vertex of Δ.

Theorem 1.2'. *Where $\boldsymbol{T}$ and $\boldsymbol{B}$ contain no problem points, there is a monotone path system Π in* $\mathbf{D}$ *which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$ if and only if*

$$(2') \qquad d(\boldsymbol{C}) \geq 0, \quad \text{for every distinguished horizontal chord } \boldsymbol{C} \text{ of } \mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}.$$

It is not difficult to see that Theorem 1.2' is equivalent with

Theorem 1.2". *Where $\boldsymbol{T}$ and $\boldsymbol{B}$ contain no problem points, there is no monotone path system Π in* $\mathbf{D}$ *which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$ if and only if $d(\boldsymbol{C}) < 0$, for some horizontal chord $\boldsymbol{C}$ of* $\mathbf{D}$.

We first stated our theorems with vertices excluded from $\boldsymbol{T}$ and $\boldsymbol{B}$ because this is one straightforward way to exclude problem points.

We will now state how to extend Theorem 1.2' to necessary and sufficient conditions for $\boldsymbol{T}$ and $\boldsymbol{B}$ to be joined by a MPS where $\boldsymbol{T}$ and $\boldsymbol{B}$ are allowed to contain problem points. Handling problem points is rather tricky and interesting.

The problem, is there a MPS in $\mathbf{D}$ pairing $\boldsymbol{T}$ with $\boldsymbol{B}$, is clearly in $\mathbb{NP}$: If there is such a MPS Π, we will show that there is one in which the total number of line segments in all the paths is bounded by a polynomial in $|\boldsymbol{T}|$ and the number of vertices of Δ. (See Remark 6.3). Given such a MPS Π, we can verify in polynomial time that Π is a MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$. It is not immediately clear that the problem is in $\mathbb{coNP}$: what can be given to enable us to verify in polynomial time that there is no such MPS? (See Figure 1.4). If there are no problem points, Theorem 1.2" provides an answer: if we are given a horizontal chord $\boldsymbol{C}$ of $\mathbf{D}$ with $d(\boldsymbol{C})<0$, we can verify this is polynomial time. But Theorem 1.3' (see below) provides an answer even if there are problem points, and thus this problem has a good characterization in the sense of Edmonds [3].

Our proofs of Theorems 1.2' and 1.3' provide good algorithms for finding a MPS if there is one

A *distinguished segment*, $\boldsymbol{K}$, of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ is a horizontal line segment of $\mathbf{D}$, whose endpoints are on Δ, whose intersection with Δ is a set of isolated points, and which contains at least one DHC. Note that $\boldsymbol{K}$ consists of a sequence of adjacent DHCs. Define

$$(3) \qquad \overline{d}(\boldsymbol{K}) := \sum\{d(\boldsymbol{C}) : \boldsymbol{C} \text{ is a horizontal chord on } \boldsymbol{K}\} + \text{the number of problem points on } \boldsymbol{K}.$$

A *maximal negative distinguished segment* or, briefly, an *n-segment*, $\boldsymbol{N}$, of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ is a distinguished segment of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ satisfying:

(i) $d(\boldsymbol{C})<0$ for every horizontal chord $\boldsymbol{C}$ on $\boldsymbol{N}$,

(ii) $\boldsymbol{N}$ is a maximal in the sense that every distinguished segment $\boldsymbol{M}$, different from $\boldsymbol{N}$ and containing $\boldsymbol{N}$, contains a horizontal chord $\boldsymbol{C}^*$ with $d(\boldsymbol{C}^*)\geq 0$.

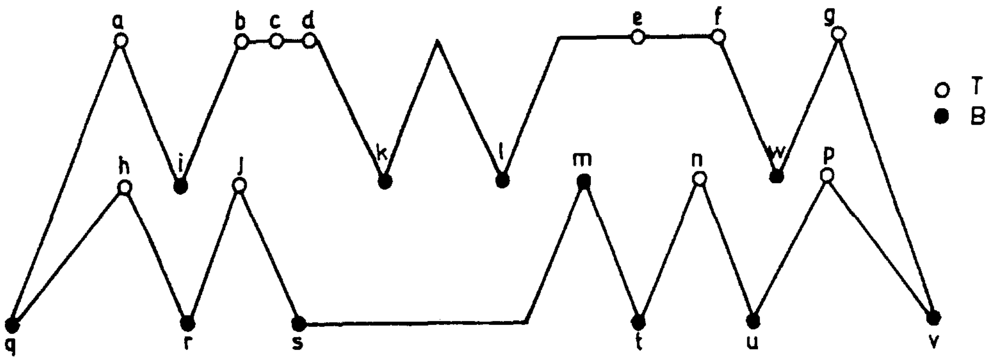

Fig. 1.4

Note that every DHC $\boldsymbol{C}$ with $d(\boldsymbol{C}) < 0$ uniquely determines an n-segment in which it is contained (and which may coincide with $\boldsymbol{C}$). Clearly, $\overline{d}(\boldsymbol{N}) \leq 1$ for every n-segment $\boldsymbol{N}$ of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$.

Theorem 1.3. *There is a monotone path system* Π *in* $\mathbf{D}$ *which pairs* $\boldsymbol{T}$ *with* $\boldsymbol{B}$ *if and only if*

$$\overline{d}(\boldsymbol{N}) \geq 0, \text{ for every } n\text{-segment } \boldsymbol{N} \text{ of } \mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}.$$

Theorem 1.3 is equivalent with

Theorem 1.3'. *There is no monotone path system* Π *in* $\mathbf{D}$ *which pairs* $\boldsymbol{T}$ *with* $\boldsymbol{B}$ *if and only if there is a distinguished segment* $\boldsymbol{K}$ *in* $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ *such that*

$$\overline{d}(\boldsymbol{K}) < 0.$$ ∎

Example. For Figure 1.4, there is no MPS which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$. Let $\boldsymbol{H} = [k,m]$. $\boldsymbol{H}$ is a distinguished segment. There are two problem points k and ℓ, on $\boldsymbol{H}$. $d(k,\ell) = -2$, $d(\ell,m) = -1$; $\overline{d}(k,\ell) = 0$; $\overline{d}(\boldsymbol{H}) = -2-1+2 < 0$. Since $d(j,k) = 1$, $d(m,n) = 0$, $\boldsymbol{H}$ is an n-segment. For $\boldsymbol{H}' = [h,j]$ and $\boldsymbol{H}'' = [n,p]$ which are also n-segments. $\overline{d}(\boldsymbol{H}') = \overline{d}(\boldsymbol{H}'') = -1-1+3 = 1$.

Presumably the theorems are true for a simple closed curve Δ in the plane, not necessarily polygonal, and for monotone paths which are not necessarily polygonal. Presumably, though, the proofs would involve technicalities which we have chosen to avoid.

Section 3 gives the proofs of Theorems 1.1' and 1.2". Section 4 prepares and Section 5 gives the proof of Theorem 1.3. In Section 6, decision and construction algorithms, solving the problem in the most general case, are described and analyzed. First, in Section 2, we give an application.

2. Hexagonal systems — an application

A *hexagonal system* **H** is a finite connected plane graph with no cut nodes in which every interior face is bounded by a regular unit hexagon. (See [6], [7], [9]). We assume that **H** is drawn in the xy-plane so that some edges are vertical (Figure 2.1a). A *perfect matching* in a graph is a set of edges which meet every node exactly once.

Let $\mathbf{D}(\mathbf{H})$ be the polygonal region bounded by the boundary of the outer face of **H**. A perfect path system of **H** is a monotone path system Π, where $\boldsymbol{T}(\Pi)$ is the set of the peaks (i.e., locally highest points) of $\mathbf{D}(\mathbf{H})$, $\boldsymbol{B}(\Pi)$ is the set of the valleys (i.e., locally lowest points) of $\mathbf{D}(\mathbf{H})$, and Π uses only edges of **H** (which are allowed to belong to the boundary of $\mathbf{D}(\mathbf{H})$. (Figure 2.1b).

Hexagonal systems **H** which have perfect path systems correspond to benzenoid hydrocarbon molecules because there is a 1–1 correspondence between perfect path systems and perfect matchings ("Kekulé structures") in **H** (Figure 2.1b); see [6]–[9].

The second author posed the question: Does every perfect path system of a hexagonal system give the same pairing of peaks with valleys? The first author's reply is the basis of this paper. The affirmative answer follows from Theorem 1.1'.

Theorem 2.1. *Let* **H** *be a hexagonal system (drawn in a fixed position) which has a perfect path system. Then every perfect path system of* **H** *gives the same pairing of the peaks with the valleys.*

Using a general result of H.-D. O. F. Gronau et al. [5] on path systems in acyclic digraphs, Theorem 2.1 enables a simple determinant formula for the number of perfect matchings of a hexagonal system to be proved in a particularly elegant way. The theorem is this:

Theorem 2.2. *Let* **H** *be a hexagonal system which has as many valleys as peaks, let p and m denote the number of perfect path systems and the number of perfect matchings of* **H**, *respectively, and let p_{ik} be the number of monotone paths in* **H** *connecting the i^{th} peak with k^{th} valley. Then*

$$m = p = |\det(p_{ik})|.$$

(See [5], [7], [8].) ∎

It is worth noting that the numbers p_{ik} can be calculated in $\boldsymbol{O}(t \cdot n)$ time by means of dynamic programming (Figure 2.1c), where t is the number of peaks and n is the number of faces of **H**.

For the hexagonal system drawn in Figures 2.1a–c we obtain

$$m = p = |\det \begin{pmatrix} 4 & 5 \\ 3 & 1 \end{pmatrix}| = 11.$$

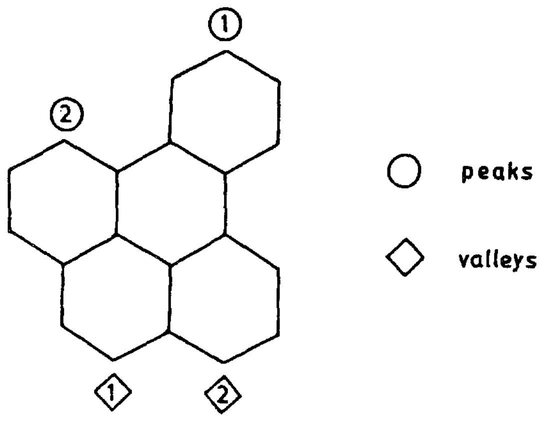

Fig. 2.1a. A hexagonal system with two peaks and two valleys.

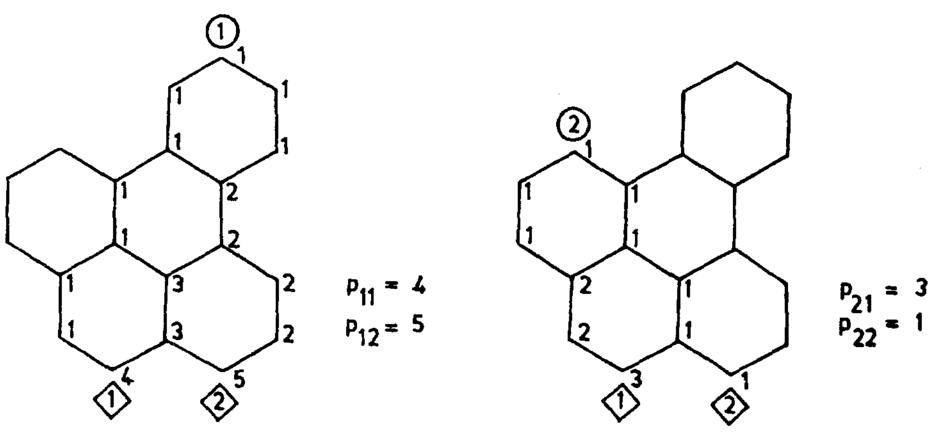

Fig. 2.1c. Calculating the numbers p_{ik}.

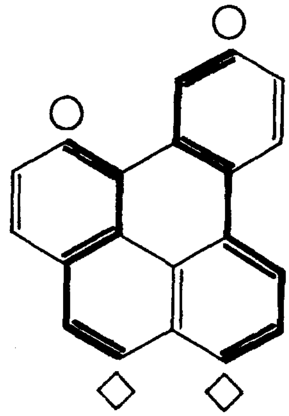

Fig. 2.1b. A perfect path system (bold face) and the corresponding perfect matching (double lines).

3. Proofs of Theorems 1.1' and 1.2'

Lemma 3.1. *Consider a monotone path π in* $\mathbf{D}$ *with top point t and bottom point b. Let $\boldsymbol{C}$ be a horizontal chord and suppose $b \in \mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C}$. Then*

(4.1) *$t \in \mathbf{L}(\boldsymbol{C})$ implies that b is an endpoint of $\boldsymbol{C}$ and a reverse valley (thus a problem point).*

(4.2) *$t \in \boldsymbol{C}$ implies that t is an endpoint of $\boldsymbol{C}$ and a reverse peak (thus a problem point).*

(4.3) *If neither t nor b is a problem point, then $t \in \mathbf{U}(\boldsymbol{C})$.*

Proof. Easy. Proof of the "only if" part of Theorem 1.2'. Suppose there are no problem points. Suppose there exists a MPS which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$. Consider a horizontal chord $\boldsymbol{C}$. By Lemma 3.1 (4.3),

(5) every $b \in \boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C})$ is a paired with a $t \in \boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C})$.

Thus

$$d(\boldsymbol{C}) = |\boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C})| - |\boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C})| \geq 0. \qquad \blacksquare$$

Corollary 3.2. *Where there are no problem points, $d(\boldsymbol{C})$ represents the number of paths of any MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ which cross $\boldsymbol{C}$.*

Proof. Recall that $\mathbf{D} - \boldsymbol{C}$ has upper component $\mathbf{U}(\boldsymbol{C})$ and lower component $\mathbf{L}(\boldsymbol{C})$. Trivially, any monotone path with its top in $\mathbf{U}(\boldsymbol{C})$ has its bottom in $\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C}$

or in $\mathbf{L}(\boldsymbol{C})$. Let Π be a MPS which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$. Using (5) it follows that precisely $d(\boldsymbol{C})$ paths of Π have their top in $\mathbf{U}(\boldsymbol{C})$ and their bottom in $\mathbf{L}(\boldsymbol{C})$; that is, precisely $d(\boldsymbol{C})$ paths cross $\boldsymbol{C}$. ∎

Recall: A distinguished horizontal chord (DHC) is a horizontal chord which contains at least one point which is either in $\boldsymbol{T} \cup \boldsymbol{B}$ or is a vertex of Δ. If we draw all the DHCs, we obtain the *canonical dissection* of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ into *canonical regions*, each of which is either a triangle with one horizontal edge or a quadrilateral with two horizontal edges and contains no point of $\boldsymbol{T} \cup \boldsymbol{B}$ in its interior.

Lemma 3.3. *Suppose there are no problem points. Suppose $d(\boldsymbol{C}) \geq 0$ for every DHC $\boldsymbol{C}$. Place $d(\boldsymbol{C})$ auxiliary points into the interior of each DHC, $\boldsymbol{C}$, distinct from any points of $\boldsymbol{T} \cup \boldsymbol{B}$ which already lie on $\boldsymbol{C}$. For canonical region $\mathbf{E}$, let $\boldsymbol{T}_{\mathbf{E}}$ be the point set consisting of those points of $\boldsymbol{T}$ and those auxiliary points which lie on the upper edge, or point (peak), of $\mathbf{E}$, and let $\boldsymbol{B}_{\mathbf{E}}$ be the point set consisting of those points of $\boldsymbol{B}$ and those auxiliary points which lie on the lower edge, or point (valley), of $\mathbf{E}$. Then $|\boldsymbol{T}_{\mathbf{E}}| = |\boldsymbol{B}_{\mathbf{E}}|$.*

Proof. Denote the upper edge, or peak, of $\mathbf{E}$ by $\boldsymbol{V}$ and the lower edge, or valley, of $\mathbf{E}$ by $\boldsymbol{W}$. Since no point of $\boldsymbol{T}$ or $\boldsymbol{B}$ is interior to $\mathbf{E}$, each point of $\boldsymbol{T} \cup \boldsymbol{B}$ is in exactly one of the sets $\boldsymbol{V}$, $\boldsymbol{W}$, $\mathbf{U}(\boldsymbol{C})$ where $\boldsymbol{C}$ is a DHC on $\boldsymbol{V}$, or $\mathbf{L}(\boldsymbol{C}')$ where $\boldsymbol{C}'$ is a DHC on $\boldsymbol{W}$. Since $|\boldsymbol{T}| = |\boldsymbol{B}|$, we have
(6)
$$\begin{aligned} &|\boldsymbol{T} \cap \boldsymbol{V}| - |\boldsymbol{B} \cap \boldsymbol{V}| + \sum\{|\boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C})| - |\boldsymbol{B} \cap \mathbf{U}(\boldsymbol{C})| : \boldsymbol{C} \text{ is a DHC on } \boldsymbol{V}\} \\ =&|\boldsymbol{B} \cap \boldsymbol{W}| - |\boldsymbol{T} \cap \boldsymbol{W}| + \sum\{|\boldsymbol{B} \cap \mathbf{L}(\boldsymbol{C}')| - |\boldsymbol{T} \cap \mathbf{L}(\boldsymbol{C}')| : \boldsymbol{C}' \text{ is a DHC on } \boldsymbol{W}\}. \end{aligned}$$

Any non-empty intersection of DHCs on $\boldsymbol{V}$ is a reverse valley. Since no point of $\boldsymbol{B} \cap \boldsymbol{V}$ is a reverse valley, each point of $\boldsymbol{B} \cap \boldsymbol{V}$ is in exactly one DHC on $\boldsymbol{V}$. Hence the l.h.s. of (6) is:

$$\begin{aligned} &|\boldsymbol{T} \cap \boldsymbol{V}| + \sum\{|\boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C})| - |\boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C})| : \boldsymbol{C} \text{ is a DHC on } \boldsymbol{V}\} \\ =&|\boldsymbol{T} \cap \boldsymbol{V}| + \sum\{d(\boldsymbol{C}) : \boldsymbol{C} \text{ is a DHC on } \boldsymbol{V}\} \\ =&|\boldsymbol{T}_{\mathbf{E}}|. \end{aligned}$$

Any non-empty intersection of DHCs on $\boldsymbol{W}$ is a reverse peak. Since no point of $\boldsymbol{T} \cap \boldsymbol{W}$ is a reverse peak, each point of $\boldsymbol{T} \cap \boldsymbol{W}$ is in exactly one DHC on $\boldsymbol{W}$. Hence the r.h.s. of (6) is:

$$\begin{aligned} &|\boldsymbol{B} \cap \boldsymbol{W}| + \sum\{|\boldsymbol{B} \cap \mathbf{L}(\boldsymbol{C}')| - |\boldsymbol{T} \cap (\mathbf{L}(\boldsymbol{C}') \cup \boldsymbol{C}')| : \boldsymbol{C}' \text{ is a DHC on } \boldsymbol{W}\} \\ =&|\boldsymbol{B} \cap \boldsymbol{W}| + \sum\{d(\boldsymbol{C}') : \boldsymbol{C}' \text{ is a DHC on } \boldsymbol{W}\} \\ =&|\boldsymbol{B}_{\mathbf{E}}|. \end{aligned}$$

Hence $|\boldsymbol{T}_{\mathbf{E}}| = |\boldsymbol{B}_{\mathbf{E}}|$. ∎

Proof of the "if" part of Theorem 1.2'. Follow the instructions of Lemma 3.3. Since for each canonical region $\mathbf{E}$, $|\boldsymbol{T}_{\mathbf{E}}| = |\boldsymbol{B}_{\mathbf{E}}|$, there exists a MPS $\Pi_{\mathbf{E}}$ in $\mathbf{E}$ with

$\boldsymbol{T}(\Pi_{\mathbf{E}}) = \boldsymbol{T}_{\mathbf{E}}$ and $\boldsymbol{B}(\Pi_{\mathbf{E}}) = \boldsymbol{B}_{\mathbf{E}}$; in fact, each path of $\Pi_{\mathbf{E}}$ can be taken to be a straight line segment, or to consist of two straight line segments (two segments are required if the points to be connected are both right corners or both left corners of the canonical domain).

Concatenating these paths at the auxiliary points, we obtain a MPS Π of $\mathbf{D}$ with $\boldsymbol{T}(\Pi)=\boldsymbol{T}$ and $\boldsymbol{B}(\Pi)=\boldsymbol{B}$.

Theorem 1.2' is now proved. ∎

We need another concept. An *interior horizontal deformation* of $\mathbf{D}$ is a homeomorphism of $\mathbf{D}$ onto itself which maps every horizontal chord of $\mathbf{D}$ onto itself. We will consider two MPSs in $\mathbf{D}$, Π_1 and Π_2 say, with $\boldsymbol{T}(\Pi_1)=\boldsymbol{T}(\Pi_2)$ and $\boldsymbol{B}(\Pi_1)=\boldsymbol{B}(\Pi_2)$ to be *equivalent* if there is an interior horizontal deformation of $\mathbf{D}$ which maps Π_1 onto Π_2.

Proof of Theorem 1.1'. Let Π be any MPS in $\mathbf{D}$ pairing $\boldsymbol{T}$ with $\boldsymbol{B}$; let $\mathbf{E}$ be any canonical region with upper edge, or point, $\boldsymbol{V}$ and lower edge, or point, $\boldsymbol{W}$. Denote the set of points which are interior to $\boldsymbol{V}$ (or $\boldsymbol{W}$) and to some path of Π (the set of the crossing points) by $\boldsymbol{V}\times\Pi$ (or $\boldsymbol{W}\times\Pi$, respectively). By Corollary 3.2 and the hypothesis that all points of $\boldsymbol{T}\cup\boldsymbol{B}$ lie on Δ, the arrangement of the points of $\boldsymbol{V}\cap(\Delta-(\boldsymbol{T}\cup\boldsymbol{B}))$, $\boldsymbol{V}\cap\boldsymbol{T}$, $\boldsymbol{V}\cap\boldsymbol{B}$, and $\boldsymbol{V}\times\Pi$ on $\boldsymbol{V}$ is (topologically) unique, i.e., the number of points of $\boldsymbol{V}\times\Pi$ in each component of $\boldsymbol{V}-(\Delta\cup\boldsymbol{T}\cup\boldsymbol{B})$ is always the same, and the same holds for $\boldsymbol{W}$. The MPS Π induces a MPS $\Pi/_{\mathbf{E}}$ in $\mathbf{E}$ pairing $(\boldsymbol{V}\cap\boldsymbol{T})\cup(\boldsymbol{V}\times\Pi)$ with $(\boldsymbol{W}\cap\boldsymbol{B})\cup(\boldsymbol{W}\times\Pi)$; clearly, $\Pi/_{\mathbf{E}}$ is unique up to equivalence, i.e. up to interior horizontal deformations of $\mathbf{E}$. This implies, by a simple induction argument, that also Π is unique up to equivalence; therefore, in particular, any MPS in $\mathbf{D}$ connecting $\boldsymbol{T}$ with $\boldsymbol{B}$ determines the same pairing of $\boldsymbol{T}$ with $\boldsymbol{B}$. ∎

Remark 3.4. When the points of $\boldsymbol{T}\cup\boldsymbol{B}$ do not all lie on Δ, we can find all possible pairings of $\boldsymbol{T}$ with $\boldsymbol{B}$ (perhaps with repetition) by varying the insertion of the auxiliary points into the DHCs in all topologically different ways, i.e., so that the number of auxiliary points in the components of $\boldsymbol{C}-(\Delta\cup\boldsymbol{T}\cup\boldsymbol{B})$ vary in all possible ways. Actually we can say more than this: By inserting the auxiliary points into the DHCs in all topologically different ways, we obtain exactly one MPS from each equivalence class. (See Section 6.)

4. Rectifiability of n-sequences — an auxiliary consideration

Before we turn to the proof of Theorem 1.3, we need some preparation.

Let $\mathbf{S}^*=(x_2,y_2,x_3,y_4,\dots,y_{2q},x_{2q+1})$ be a sequence of integers where $x_{2\lambda+1}\in\{0,1\}$ $(\lambda=0,1,\dots,q;q\geq 0)$. Put $\{1,3,5,\dots,2q+1\}=:\boldsymbol{I}=\boldsymbol{I}_0\cup\boldsymbol{I}_1$ where $2\lambda+1\in\boldsymbol{I}_k$ if and only if $x_{2\lambda+1}=k$ $(k=0,1)$. (Note that one of $\boldsymbol{I}_0$, $\boldsymbol{I}_1$ may be empty.) If $x_1=1$, add a first element y_0 to $\mathbf{S}^*$, if $x_{2q+1}=1$, add a last element y_{2q+2} to $\mathbf{S}^*$, where y_0 and/or y_{2q+2} are assumed to be non-negative integers (the values of y_0 and y_{2q+2} will be fixed in the applications; here, they are immaterial); thus $\mathbf{S}^*$ is turned into a sequence $\mathbf{S}$. Put $\boldsymbol{J}^*:=\{2,4,\dots,2q\}$ and let $\boldsymbol{J}$ denote the set of all even indices occurring in $\mathbf{S}$; put $\partial\boldsymbol{J}:=\boldsymbol{J}-\boldsymbol{J}^*$ ($\partial\boldsymbol{J}$ may be empty).

A *shift function* is a function f, defines on $\boldsymbol{I}_1$, which takes each of the odd indices from $\boldsymbol{I}_1$ to an adjacent even index (i.e., $f(i)\in\{i-1,i+1\}$, $i\in\boldsymbol{I}_1$).

Put

$$\text{(7)} \qquad \begin{cases} x_i' = 0, & i \in \boldsymbol{I} \\ y_j' = y_j + \sum\{x_k : k \in f^{-1}(j)\}, & j \in \boldsymbol{J}. \end{cases}$$

(That is, each $x_i \neq 0$ is "shifted" to a neighbouring y, either to its left or to its right, according as $f(i) = i-1$ or $i+1$, and added to it.)

By this operation, $\mathbf{S}$ is transformed into a new sequence $\mathbf{S}'$ where $\boldsymbol{I}_1' = \emptyset$. Note that the sum of all elements remains unchanged:

$$\text{(8)} \qquad \sum_{i \in \boldsymbol{I}} x_i + \sum_{j \in \boldsymbol{J}} y_j = \sum_{j \in \boldsymbol{J}} y_j';$$

further,

$$\text{(9)} \qquad y_j \le y_j' \le y_j + 2, \quad j \in \boldsymbol{J}^*$$

and, if y_0 and/or y_{2q+2} are present,

$$\text{(10)} \qquad y_0 \le y_0' \le y_0 + 1, \quad y_{2q+2} \le y_{2q+2}' \le y_{2q+2} + 1.$$

$\mathbf{S}$ is called *rectifiable* if there is a shift function f such that

$$\text{(11)} \qquad y_j' \ge 0, \text{ for all } j \in \boldsymbol{J}.$$

$\mathbf{S}$ is called an *n-sequence* if $y_j < 0$, for all $j \in \boldsymbol{J}^*$.

Next we shall derive a simple necessary and sufficient condition for an n-sequence to be rectifiable, determine explicitly the set $\mathcal{R}$ of all rectifiable n-sequences, and list all shift functions which rectify a given n-sequence from $\mathcal{R}$.

Let $\mathbf{S}$ be a sequence which is rectifiable. Consider any shift function which rectifies $\mathbf{S}$. From (8), (10), and (11) we obtain

$$\begin{aligned} \sum_{i \in \boldsymbol{I}} x_i + \sum_{j \in \boldsymbol{J}^*} y_j &= \sum_{i \in \boldsymbol{I}} x_i + \sum_{j \in \boldsymbol{J}} y_j - \sum_{j \in \partial \boldsymbol{J}} y_j \\ &= \sum_{j \in \boldsymbol{J}} y_j' - \sum_{j \in \partial \boldsymbol{J}} y_j = \sum_{j \in \boldsymbol{J}^*} y_j' + \sum_{j \in \partial \boldsymbol{J}} (y_j' - y_j) \\ &\ge 0, \end{aligned}$$

that is,
(12) the sum of all elements in $\mathbf{S}^*$ is non-negative,
or more briefly,

$$\text{(12}'\text{)} \qquad |\boldsymbol{I}_1| + \sum_{j \in \boldsymbol{J}^*} y_j \ge 0.$$

We shall show that for on n-sequence $\mathbf{S}$, condition (12) (or, equivalently, (12′)) is not only necessary but also sufficient for $\mathbf{S}$ to be rectifiable.

Let $\mathbf{S}$ be an n-sequence which satisfies (12). Clearly,

$$|\boldsymbol{I}_1| \le q+1, \quad \sum_{j\in \boldsymbol{J}^*} y_j \ge -q,$$

thus

$$|\boldsymbol{I}_1| + \sum_{j\in \boldsymbol{J}^*} y_j \le 1.$$

From (12′) we obtain that one of the following cases holds.

Case 1. $|\boldsymbol{I}_1| = q, \quad \sum_{j\in \boldsymbol{J}^*} y_j = -q.$

This means that exactly one of the x_i — say, x_{i_0} — is zero and all others are equal to 1, and all y_j $(j \in \boldsymbol{J}^*)$ are equal to –1.

There is a unique shift function rectifying $\mathbf{S}$, namely,

$$f(i) = \begin{cases} i+1, & i = 1, 3, \ldots, i_0 - 2 \\ i-1, & i = i_0+2, i_0+4, \ldots, 2q+1. \end{cases}$$

Case 2. $|\boldsymbol{I}_1| = q+1, \quad \sum_{j\in \boldsymbol{J}^*} y_j = -(q+1).$

This means that all x_i are equal to 1, and exactly one of the $y_j (j \in \boldsymbol{J}^*)$ — say, y_{j_0} — is equal to –2, and all others are equal to –1.

There is a unique shift function rectifying $\mathbf{S}$, namely,

$$f(i) = \begin{cases} i+1, & i = 1, 3, \ldots, j_0 - 1 \\ i-1, & i = j_0+1, j_0+3, \ldots, 2q+1. \end{cases}$$

Case 3. $|\boldsymbol{I}_1| = q+1, \sum_{j\in \boldsymbol{J}^*} y_i = -q.$

This means that all x_i are equal to 1, and all $y_j (j \in \boldsymbol{J}^*)$ are equal to –1.

In this case, there are exactly $q+2$ shift functions rectifying $\mathbf{S}$, one for each $j \in \boldsymbol{J}$, namely,

$$f_j(i) = \begin{cases} i+1, & i = 1, 3, \ldots, j - 1 \\ i-1, & i = j+1, j+3, \ldots, 2q+1 \end{cases}$$

(see Figure 4.1).

Thus we have proved

Lemma 4.1. *An n-sequence $\mathbf{S}$ is rectifiable if and only if*
(12) *the sum of all elements in $\mathbf{S}^*$ is non-negative,*
that is, if and only if

(12′)
$$|\boldsymbol{I}_1| + \sum_{j\in \boldsymbol{J}^*} y_j \ge 0.$$

All rectifiable n-sequences and the corresponding shift functions rectifying them are explicitly described under Cases 1, 2, 3 above. ∎

$j=0$

index:	0	1	2	3	4	5	…	2q	2q+1	2q+2
S=	$(y_0,$	1,	−1,	1,	−1,	1,	…,	−1,	1,	$y_{2q+2})$
	⟵	⟵	⟵	⟵	⟵	⟵		⟵	⟵	
S′=	$(y_0+1,$	0,	0,	0,	0,	0,	…,	0,	0,	$y_{2q+2})$

$0<j<2q+2$

index:	0	1	2	…	j−1	j	j+1	…	2q	2q+1	2q+2
S=	$(y_0,$	1,	−1,	…,	1,	−1,	1,	…,	−1,	1,	$y_{2q+2})$
		⟶	⟶		⟶	⟶⟵	⟵		⟵	⟵	
S′=	$(y_0,$	0,	0,	…,	0,	1,	0,	…,	0,	0,	$y_{2q+2})$

$j=2q+2$

index:	0	1	2	3	4	…	2q−1	2q	2q+1	2q+2
S=	$(y_0,$	1,	−1,	1,	−1,	…,	1,	−1,	1,	$y_{2q+2})$
		⟶	⟶	⟶	⟶	…	⟶	⟶	⟶	⟶
S′=	$(y_0,$	0,	0,	0,	0,	…,	0,	0,	0,	$y_{2q+2}+1)$

Fig. 4.1

5. Proof of Theorem 1.3

A reverse peak or a reverse valley, p, lies on two DHCs, one to its left (call it $\boldsymbol{C}_\ell$) and one to its right (call it $\boldsymbol{C}_r$). Such a point p also lies in three canonical regions, one of them lies to the left of p (call it $\mathbf{E}_\ell$), one lies to the right of p (call it $\mathbf{E}_r$), and the third (call it $\mathbf{E}'$) lies above or below p depending on whether p is a reverse peak or a reverse valley. See Figure 5.1.

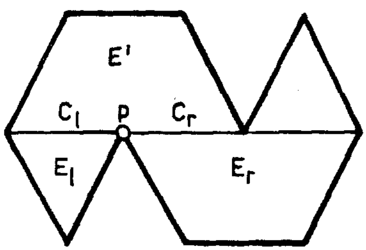

Fig. 5.1a. The elementary regions containing reverse peak p.

Recall that a problem point is a reverse peak in $\boldsymbol{T}$ or a reverse valley in $\boldsymbol{B}$.

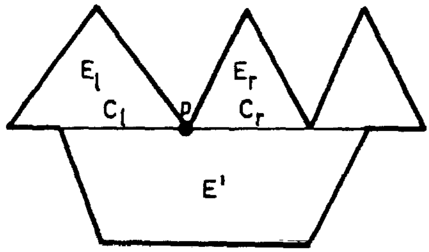

Fig. 5.1b., The elementary regions containing reverse valley p.

Lemma 5.1. *There is a MPS in* $\mathbf{D}$ *which pairs* $\boldsymbol{T}$ *with* $\boldsymbol{B}$ *if and only if there is a (small) shift of each problem point to the left or to the right such that after all shifts are made,* $d(\boldsymbol{C}) \geq 0$ *for every DHC* $\boldsymbol{C}$.

Proof. Let p be a problem point. First, suppose $p \in \boldsymbol{T}$. Let π be a monotone path with p as its top point. $\pi - p$ intersects exactly one of $\mathbf{E}_\ell$ and $\mathbf{E}_r$. The same is true if $p \in \boldsymbol{B}$ and π is a monotone path with p as its bottom point. So in either case, π is almost the same as a monotone path π' where p has been shifted ($p \to p'$) left or right depending on whether $\pi - p$ intersects $\mathbf{E}_\ell$ or $\mathbf{E}_r$ (see Figure 5.2). Once a problem point has been shifted, it is no longer a problem point. Lemma 5.1 now follows from Theorem 1.2'.

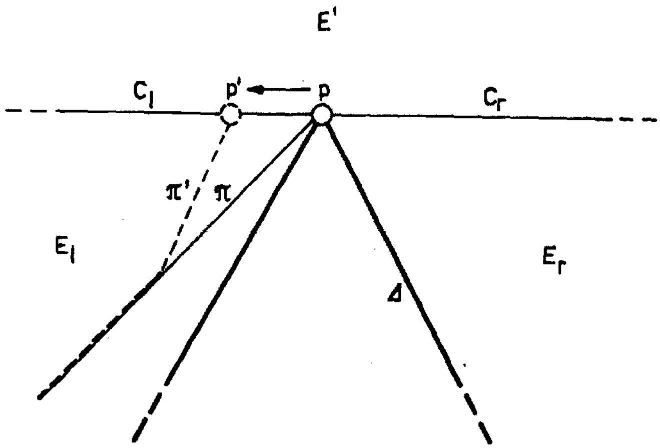

Fig. 5.2

Let $\boldsymbol{N}$ be an n-segment of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. Recall that this means that $\boldsymbol{N}$ is a horizontal line segment in $\mathbf{D}$ which consists of a sequence of consecutive DHCs, say $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_q$ in a left-to-right order, with $d(\boldsymbol{C}_\lambda) < 0$ $(\lambda = 1, 2, \ldots, q; q \geq 1)$ and that $\boldsymbol{N}$ is inclusion-maximal with respect to these properties. Let $[a, b]$ denote a horizontal line segment with left endpoint a and right endpoint b. Assume $\boldsymbol{C}_\lambda = [p_{\lambda-1}, p_\lambda]$; clearly, $\boldsymbol{N} = [p_0, p_q]$, $\boldsymbol{C}_\lambda \cap \boldsymbol{C}_{\lambda+1} = \{p_\lambda\}$ $(\lambda = 1, 2, \ldots, q-1)$, and $\boldsymbol{N} \cap \Delta = \{p_0, p_1, \ldots, p_q\}$. $\boldsymbol{N}$ determines a sequence

$$\mathbf{S}^*(\boldsymbol{N}) = \{x_1, y_2, x_3, y_4, \ldots, y_{2q}, x_{2q+1}\}$$

of integers as follows:

$$\left\{\begin{array}{l} x_{2\lambda+1} := \begin{cases} 1 & \text{if } p_\lambda \text{ is a problem point,} \\ 0 & \text{if } p_\lambda \text{ is not a problem point,} \end{cases} \\ y_{2\mu} := d(\boldsymbol{C}_\mu) \\ (\lambda = 0, 1, \ldots, q; \quad \mu = 1, 2, \ldots, q). \end{array}\right. \tag{13}$$

If $x_1 = 1$ $(x_{2q+1} = 1)$ then $p_0(p_q)$ is a problem point and there is in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ a DHC with right endpoint p_0 (left endpoint p_q): call it $\boldsymbol{C}_0$ $(\boldsymbol{C}_{q+1})$. By the maximality property of $\boldsymbol{N}$, $d(\boldsymbol{C}_0) \geq 0$, $d(\boldsymbol{C}_{q+1}) \geq 0$ (if $\boldsymbol{C}_0$ or $\boldsymbol{C}_{q+1}$, respectively, exists). If $x_1 = 1$, add $y_0 := d(\boldsymbol{C}_0)$ as a first element to the sequence $\mathbf{S}^*(\boldsymbol{N})$; if $x_{2q+1} = 1$, add $y_{2q+2} := d(\boldsymbol{C}_{q+1})$ as a last element to the sequence $\mathbf{S}^*(\boldsymbol{N})$ thus turning $\mathbf{S}^*(\boldsymbol{N})$ into a sequence $\mathbf{S}(\boldsymbol{N})$. Clearly, $\mathbf{S}(\boldsymbol{N})$ is an n-sequence.

We will now study the effect of shifting problem points lying on $\boldsymbol{N}$ on the sequence $\mathbf{S}(\boldsymbol{N})$.

Lemma 5.2. *By shifting a problem point to the left,*
$d(\boldsymbol{C}_\ell)$ stays the same,
$d(\boldsymbol{C}_r)$ increases by one,
$d(\boldsymbol{C})$ stays the same for any horizontal chord $\boldsymbol{C} \neq \boldsymbol{C}_\ell, \boldsymbol{C}_r$.

Note that Lemma 5.2 says that shifting a problem point changes $d(\boldsymbol{C})$ only for the chord that the point is shifted *off*.

Proof of Lemma 5.2. It suffices to prove the lemma for a problem point which is a reverse valley $p \in \boldsymbol{B}$. Moving p to the left onto $\boldsymbol{C}_\ell$ does not affect $d(\boldsymbol{C}_\ell)$ because p was already on $\boldsymbol{C}_\ell$. However, this does take p off $\boldsymbol{C}_r$ and into the lower component $\mathbf{L}(\boldsymbol{C}_r)$ of $\mathbf{D} - \boldsymbol{C}_r$. Then since $d(\boldsymbol{C}_r) = |\boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C}_r)| - |\boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}_r) \cup \boldsymbol{C}_r)|$, where $\mathbf{U}(\boldsymbol{C}_r)$ is the upper component of $\mathbf{D} - \boldsymbol{C}_r$, $d(\boldsymbol{C}_r)$ increases by one, because $|\boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}_r) \cup \boldsymbol{C}_r)|$ decreases by one. Now consider any horizontal chord $\boldsymbol{C} \neq \boldsymbol{C}_\ell$, $\boldsymbol{C}_r$. $\boldsymbol{C}_\ell$ and $\boldsymbol{C}_r$ both lie either in $\mathbf{U}(\boldsymbol{C})$ or in $\mathbf{L}(\boldsymbol{C})$, thus shifting p does not affect $d(\boldsymbol{C})$. ∎

Lemma 5.3. *Let $\boldsymbol{N}$ be an n-segment of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. The problem points lying on $\boldsymbol{N}$ can be shifted on $\boldsymbol{N}$ in such a way that, as the result of the shifting process, $d(\boldsymbol{C}) \geq 0$ for all DHCs $\boldsymbol{C}$ on $\boldsymbol{N}$* (i) *if and* (ii) *only if the corresponding n-sequence $\mathbf{S}(\boldsymbol{N})$ is rectifiable.*

Proof. (i) Suppose that $\mathbf{S}(\boldsymbol{N})$ is rectifiable; let f be any shift function rectifying $\mathbf{S}(\boldsymbol{N})$. For $\lambda \in \{0, 1, \ldots, q\}$, p_λ is a problem point on $\boldsymbol{N}$ if and only if $x_{2\lambda+1} = 1$, that is, if and only if $2\lambda+1 \in \boldsymbol{I}_1$. Let $i = 2\lambda+1 \in \boldsymbol{I}_1$: shift p_λ to the left if $f(i) = i+1$

and to the right if $f(i)=i-1$. By Lemma 5.2, after the shift is made, $d(\boldsymbol{C}_\mu)=y'_{2\mu}(2\mu\in\boldsymbol{J})$ and as the effect of the rectification, $y'_{2\mu}\geq 0$, thus $d(\boldsymbol{C})\geq 0$ for every DHC $\boldsymbol{C}$ on $\boldsymbol{N}$.

(ii) is proved by reversing the proof of (i). ∎

Remark 5.4 The proof of Lemma 5.3 shows that there is a 1–1 correspondence between the set of shifts of problem points which make all numbers $d(\boldsymbol{C})$ ($\boldsymbol{C}$ on $\boldsymbol{N}$) non-negative and the set of shift functions which rectify $\mathbf{S}(\boldsymbol{N})$. Now Lemma 4.1 gives us all information we need in order to find, in linear time, for any given n-segment $\boldsymbol{N}$ all possible shifts of the problem points which result in non-negative numbers $d(\boldsymbol{C})$.

It is now almost clear how to complete the proof of Theorem 1.3. By Lemma 5.1, a MPS in $\mathbf{D}$ pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ exists if and only if the problem points can be shifted in such a way that no n-segment remains (in particular, if there are no problem points, then such a MPS exists if and only if there is no n-segment in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$; this is in accordance with Theorem 1.2'). According to Lemma 5.3, this is the case if and only if for every n-segment $\boldsymbol{N}$ in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, the corresponding n-sequence $\mathbf{S}(\boldsymbol{N})$ is rectifiable. By Lemma 4.1, this takes place if and only if, for every n-segment $\boldsymbol{N}$ in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, the sum of all elements in $\mathbf{S}^*(\boldsymbol{N})$ — call it $s^*(\boldsymbol{N})$ — is non-negative. Now, $s^*(\boldsymbol{N})=|\boldsymbol{I}_1|+\sum_{j\in\boldsymbol{J}^*} y_j$ (see ($12'$)) where $|\boldsymbol{I}_1|$ is the number of problem points on $\boldsymbol{N}$, $y_j=y_{2\mu}$ is equal to $d(\boldsymbol{C}_\mu)$ (see (13)), and $\sum_{j\in\boldsymbol{J}^*}=\sum_{\mu=1}^{q} y_{2\mu}=\sum\{d(\boldsymbol{C}):\boldsymbol{C}$ is a horizontal chord on $\boldsymbol{N}\}$, thus $s^*(\boldsymbol{N})=\overline{d}(\boldsymbol{N})$ (see (3)). It follows that there is a MPS in $\mathbf{D}$ which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$ if and only if $\overline{d}(\boldsymbol{N})\geq 0$ for every n-segment $\boldsymbol{N}$ of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ — this is precisely what Theorem 1.3 says. ∎

6. Computing one or all monotone path systems

First, we will describe the algorithm in general terms. Then we will describe how to compute the numbers $d(\boldsymbol{C})$. Finally, we give further implementation details and analyze the complexity of the algorithm.

A. The algorithm

Given $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, a polygonal region $\mathbf{D}$ with point sets $\boldsymbol{T}$ and $\boldsymbol{B}$ inserted, we now describe our algorithm to find a MPS which pairs $\boldsymbol{T}$ with $\boldsymbol{B}$ or determine that there isn't one. In square parentheses we give the modification of the algorithm which will find all MPSs pairing $\boldsymbol{T}$ with $\boldsymbol{B}$, i.e., precisely one MPS from each equivalence class.

(i) Find the canonical dissection and list all the canonical regions of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. Make a list Λ of all DHCs of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. Calculate $d(\boldsymbol{C})$ for each DHC $\boldsymbol{C}$. Determine all problem points of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. Use Λ to find and list all n-segments of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$. For each n-segment $\boldsymbol{N}$ of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, calculate

$$\overline{d}(\boldsymbol{N}) = \sum\{d(\boldsymbol{C}) : \boldsymbol{C} \text{ is a DHC on } \boldsymbol{N}\} + \text{number of problem points on } \boldsymbol{N}.$$

By Theorem 1.3, a MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ exists if and only if

$$\overline{d}(\boldsymbol{N}) \geq 0 \text{ for all } n\text{-segments } \boldsymbol{N} \text{ of } \mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}. \tag{14}$$

If (14) is not satisfied, the algorithm stops. Suppose that (14) is satisfied.

(ii) There may be problem points which do not lie on an n-segment: such problem points are those for which $d(\boldsymbol{C}_\ell) \geq 0$ and $d(\boldsymbol{C}_r) \geq 0$; they may be called "non-essential". (Note that non-essential problem point p may be considered as defining a degenerate n-segment $\boldsymbol{N}_p$ with $q=0$ and $d(\boldsymbol{N}_p)=1$.)

(iii) Perform shifts [all possible shifts if you want to find all MPSs] of all problem points as described in Section 5, where the non-essential problem points may be shifted arbitrarily to the left or to the right, such that when all problem points have been shifted a new configuration $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ is obtained in which there are no n-segments (see Section 5); denote the modified numbers $d(\boldsymbol{C})$ by $d'(\boldsymbol{C})$ (which are all non-negative). Consider the set $\boldsymbol{C}'$ of DHCs of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ and insert $d'(\boldsymbol{C})$ auxiliary points into each $\boldsymbol{C} \in \boldsymbol{C}'$ (as described in Section 3, Lemma 3.3, and the proof of the "if" part of Theorem 1.2'). [If you want to find all MPSs pairing $\boldsymbol{T}$ with $\boldsymbol{B}$, then for every such $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, consider the set $\boldsymbol{C}'$ of all DHCs and insert in all topologically different ways $d'(\boldsymbol{C})$ auxiliary points into each $\boldsymbol{C} \in \boldsymbol{C}'$.] Note that no auxiliary point may be placed between the original and the shifted position of a problem point.

(iv) For the configuration [all configurations if you want to find all MPSs] so obtained, construct a MPS pairing $\boldsymbol{T}'$ with $\boldsymbol{B}'$ (see Section 3, proof of the "if" part of Theorem 1.2'.) Then let the problem points return to their original positions and adjust the MPS accordingly. You now have a MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ [precisely one MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ from each equivalence class.]

B. Computing the numbers $d(C)$

Associate with $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ a directed graph $\mathbf{G} = \mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}) = (\boldsymbol{N}, \boldsymbol{A})$ with node-set $\boldsymbol{N}$ and arc-set $\boldsymbol{A}$ in the following way:

(a) the set $\boldsymbol{N}$ of nodes of $\mathbf{G}$ is in 1–1 correspondence with the set $\mathbf{E}$ of canonical regions of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$;

(b) there is an arc with tail node n' and head node n'' in $\mathbf{G}$ if and only if the corresponding regions $\mathbf{E}'$ and $\mathbf{E}''$ are adjacent in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ and $\mathbf{E}'$ lies above $\mathbf{E}''$, i.e., if and only if $\mathbf{E}'$ and $\mathbf{E}''$ have a DHC in common which is on the lower edge of $\mathbf{E}'$ and the upper edge $\mathbf{E}''$.

Note that the set $\boldsymbol{A}$ of set of arcs of $\mathbf{G}$ is in 1–1 correspondence with the set $\boldsymbol{C}'$ of DHCs in $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$.

$\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ can be considered a planar map. If we ignore the directions of $\mathbf{G} = \mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}})$, we obtain what is known in as the "weak planar dual" of the planar map $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ (weak because no node is assigned to the outer face; i.e. the region exterior to $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$).

Clearly, $\mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}})$ is connected, and it is well-known (see for example [1]) that $\mathbf{G}$ is a tree. (This is easily seen: since the deletion of any DHC of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$ disconnects $\mathbf{D}$, the deletion of any arc of $\mathbf{G}$ disconnects $\mathbf{G}$.)

For canonical region $\mathbf{E}$ with upper edge or point $\boldsymbol{V}$ and lower edge or point $\boldsymbol{W}$, define

$$\delta(\mathbf{E}) := |\boldsymbol{T} \cap \boldsymbol{V}| - |\boldsymbol{B} \cap \boldsymbol{W}|. \tag{15}$$

Consider $\mathbf{G} = \mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}) = (\boldsymbol{N}, \boldsymbol{A})$. Let node $n \in \boldsymbol{N}$ correspond to canonical region $\mathbf{E}$. Define *charge* of node n to be

$$q(n) := \delta(\mathbf{E}) \tag{16}$$

and the *charge* of subgraph $\mathbf{G}'$ of $\mathbf{G}$ to be

$$q(\mathbf{G}') := \sum \{q(n) : n \text{ is a node of } \mathbf{G}'\}. \tag{17}$$

Note that

$$q(\mathbf{G}) = \sum_{n \in \boldsymbol{N}} q(n) = \sum_{\mathbf{E}} \delta(\mathbf{E}) = |\boldsymbol{T}| - |\boldsymbol{B}| = 0. \tag{18}$$

Let arc $a \in \boldsymbol{A}$ correspond to DHC $\boldsymbol{C} \in \boldsymbol{C}'$. The graph $\mathbf{G} - a$ consists of two branches of $\mathbf{G}$: an upper branch $\mathbf{G_U}(a)$ which corresponds to $\mathbf{G}(\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C})$ and a lower branch $\mathbf{G_L}(a)$ which corresponds to $\mathbf{G}(\mathbf{L}(\boldsymbol{C}) \cup \boldsymbol{C})$. By (18),

$$q(\mathbf{G_U}(a)) = -q(\mathbf{G_L}(a)),$$

and by (17), (16), and (15),

$$\begin{aligned} q(\mathbf{G_U}(a)) &:= \sum \{q(n) : n \text{ is a node of } \mathbf{G_U}(a)\} \\ &= \sum \{\delta(\mathbf{E}) : \mathbf{E} \text{ is a canonical region of } \mathbf{D}^{\boldsymbol{T},\boldsymbol{B}} \text{ contained in } \mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C}\} \\ &= |\boldsymbol{T} \cap \mathbf{U}(\boldsymbol{C})| - |\boldsymbol{B} \cap (\mathbf{U}(\boldsymbol{C}) \cup \boldsymbol{C})| = d(\boldsymbol{C}). \end{aligned}$$

Thus the numbers $d(\boldsymbol{C})$ can be easily computed by traversing the tree $\mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}})$ starting from the sources.

C. Implementation and analysis of the algorithm

The number of MPSs pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ can be exponential in $|\boldsymbol{T}|$: The configuration of Figure 1.0 has two pairings. Thus the configuration of Figure 6.1 has $2^{|\boldsymbol{T}|/2}$ pairings.

Thus, although our algorithm to find all MPSs pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ is clearly polynomial in the size of the output, it can be exponential in the size of the input.

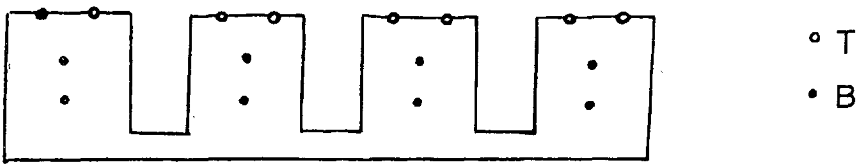

Fig. 6.1. A configuration with $2^{|T|/2}$ *pairings.*

Theorem 6.1. *Where input polygon* Δ *has* n *vertices and* $|\boldsymbol{T}| = |\boldsymbol{B}| = m$, *our algorithm to find a MPS which pairs* $\boldsymbol{T}$ *with* $\boldsymbol{B}$ *can be implemented in* $O((n+m)\max\{m, \log(n+m)\})$ *time.*

Proof. We will assume polygon Δ is given by listing its vertices in clockwise order.

If we draw all the horizontal chords of the polygonal region $\mathbf{D}$ determined by polygon Δ which have at least one vertex of Δ as an endpoint, we obtain what is known in the literature as the *trapezoidization* of $\mathbf{D}$, $\mathrm{Tr}(\mathbf{D})$. The associated elementary regions are called trapezoids in the literature, although, of course, they may be either trapezoids or triangles. $\mathrm{Tr}(\mathbf{D})$ can be obtained in $O(n \log n)$ time [2], [4]. To obtain the canonical dissection of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, we will start with $\mathrm{Tr}(\mathbf{D})$ and then split the elementary regions by drawing horizontal chords through the points of $\boldsymbol{T} \cup \boldsymbol{B}$ which are not vertices of Δ.

The algorithms [2], [4] which find $\mathrm{Tr}(\mathbf{D})$ also find $\mathbf{G}(\mathrm{Tr}(\mathbf{D}))$ and, without increasing the complexity, can be made to output the following for $\mathrm{Tr}(\mathbf{D})$:

(a) A list for the elementary regions.
(b) For each elementary region, its three or four corner points.
(c) A list of the DHCs, each of which is doubly linked to the elementary regions containing it and the vertices it contains.

Starting with $\mathrm{Tr}(\mathbf{D})$, $\mathbf{G}(\mathrm{Tr}(\mathbf{D}))$, and the above lists for $\mathrm{Tr}(\mathbf{D})$, we obtain the canonical dissection of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, $\mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}})$, and the above lists for the canonical dissection of $\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}}$, where in (c), "vertices of Δ" is replaced by "vertices of Δ and points of $\boldsymbol{T} \cup \boldsymbol{B}$". For each point $p \in \boldsymbol{T} \cup \boldsymbol{B}$ which is not a vertex of Δ, determine which region $\mathbf{E}$ of the current dissection p lies in. Draw a horizontal chord of $\mathbf{E}$ through p to split $\mathbf{E}$ into two regions, an upper region $\mathbf{E}_{\mathrm{U}}$ and a lower region $\mathbf{E}_{\mathrm{L}}$. Add $\mathbf{E}_{\mathrm{U}}$ and $\mathbf{E}_{\mathrm{L}}$ to the list of regions and the chord to the list of DHCs. To update the graph of the current dissection, split the node n corresponding to $\mathbf{E}$ into two nodes n_{U} corresponding to $\mathbf{E}_{\mathrm{U}}$ and n_{L} corresponding to $\mathbf{E}_{\mathrm{L}}$. All arcs which had n as their head should now have n_{U} as their head. All arcs which had n as their tail should now have n_{L} as their tail. Also add an arc from n_{U} to n_{L}.

Claim 6.2. *In any canonical dissection, there are at most* $2m+2n-2$ *DHCs and at most* $2m+2n-1$ *elementary regions.*

Proof of Claim 6.2 A vertex of Δ lies on at most two DHCs. A point of $\boldsymbol{T} \cup \boldsymbol{B}$ which is not a vertex lies on one DHC. The vertex of Δ or point of $\boldsymbol{T} \cup \boldsymbol{B}$ which has the largest y-coordinate and the one with the smallest y-coordinate do not lie on any DHC. ∎

By Claim 6.2 we can determine which region a point lies in $O(2m+2n-1)$ time, and do the rest of the work described above for that point in constant time. Doing this for all points of $\boldsymbol{T} \cup \boldsymbol{B}$ which are not vertices thus takes at most $O(2m(2m+2n-1))$ time.

Now we will determine which of the points of $\boldsymbol{T} \cup \boldsymbol{B}$ are problem points. Suppose that the vertices of Δ given in clockwise order are $v_i = (x_i, y_i), 1 \le i \le n$.

$v_i \in \boldsymbol{T}$ is a problem point $\Leftrightarrow (y_i > y_{i-1})$ and $(y_i > y_{i+1})$ and (v_i lies on two DHCs).
$v_i \in \boldsymbol{B}$ is a problem point $\Leftrightarrow (y_i < y_{i-1})$ and $(y_i < y_{i+1})$ and (v_i lies on two DHCs).

(19) Since a problem point is both a vertex and a member of $\boldsymbol{T} \cup \boldsymbol{B}$, the number of problem points is less than or equal to $\max\{n, 2m\}$.

Thus in time $O(\max\{n, 2m\})$, we can determine which points are problem points.

For each elementary region $\mathbf{E}$ with upper edge or point $\boldsymbol{V}$ and lower edge or point $\boldsymbol{W}$, calculate $\delta(\mathbf{E}) := |\boldsymbol{T} \cap \boldsymbol{V}| - |\boldsymbol{B} \cap \boldsymbol{W}|$. Then calculate the numbers $d(\boldsymbol{C})$ by traversing $\mathbf{G}(\mathbf{D}^{\boldsymbol{T},\boldsymbol{B}})$ as described in Section B above. This can be done in $O(2m+2n-1)$ time.

n-segments can be found in a straightforward way as follows. Consider, for each DHC $\boldsymbol{C}$ with $d(\boldsymbol{C}) < 0$, an ordered pair consisting of the y-coordinate of $\boldsymbol{C}$ and the x-coordinate of the left endpoint of $\boldsymbol{C}$. Sort these pairs lexicographically. Then read the sublist with the same y-coordinate in order, checking if the x-coordinates of the right endpoint of the current DHC and the left endpoint of the next DHC are the same. This can be done in $O((2m+2n-2)\log(2n+2m-2))$ time. Then, the numbers $\overline{d}(\boldsymbol{N})$ for n-segments $\boldsymbol{N}$ can be calculated in $O(2m+2n-2)$ time.

Let k be the number of problem points. By (19), $k \le \max\{n, 2m\}$. A single shift of each problem point an the new numbers $d'(\boldsymbol{C})$ can be calculated in $O(k)$ time. Clearly, for every DHC $\boldsymbol{C}$, $d'(\boldsymbol{C}) \le m$. Thus inserting $d'(\boldsymbol{C})$ auxiliary points into each $\boldsymbol{C} \in \boldsymbol{C}'$ can be done in $O(m(2m+2n-2))$ time.

To draw a monotone path in an elementary region $\mathbf{E}$ requires two straight line segments if the points to be connected are both right corners or both left corners of the region and one straight line segment otherwise. The number of paths of a MPS $\Pi_{\mathbf{E}}$ in $\mathbf{E}$ is at most m, so the number of line segments required for $\Pi_{\mathbf{E}}$ is at most $m+2$. Thus we can construct a MPS $\Pi_{\mathbf{E}}$ for each of the elementary regions $\mathbf{E}$ with a total of at most $(2m+2n-1)(m+2)$ line segments.

Remark 6.3 This is a MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ with a total of at most $(2m+2n-1)(m+2)$ line segments. This can be constructed in $O((2m+2n-1)(m+2))$ time. Returning problem points to their original positions can be done in $O(k)$ time.

It follows that the overall complexity of the algorithm for finding a MPS pairing $\boldsymbol{T}$ with $\boldsymbol{B}$ is $O((n+m)\max\{m, \log(n+m)\})$. ∎

References

[1] B. Chazelle: A theorem on polygon cutting with applications, *Proc. 23rd Annual IEEE Symp. Found. Comput. Sci.*, 1982, 67–72.

[2] B. Chazelle, and J. Incerpi: Triangulation and shape complexity, *ACM Trans. on Graphics* **3** (1984), 135–152.

[3] J. Edmonds: Minimum partition of a matroid into independent subsets, Proceedings of the First Conference on Matroids, *J. Res. Nat. Bur. Standards Sect. B* **69** (1965) 67–72.

[4] A. Fournier, and D. Y. Montuno: Triangulating simple polygons and equivalent problems, *ACM Trans. on Graphics* **3** (1984), 153–174.

[5] H.-D. O. F. Gronau, W. Just, W. Schade, P. Scheffer, and J. Wojciechowski: Path systems in acyclic directed graphs, Proceedings of the International Conference "Combinatorial Analysis and Its Applications" (Pokrzywna, Poland, Sept. 16–20, 1985), *Zastosowania Matematyki* **19** (4), to appear.

[6] P. John, and H. Sachs: Calculating the number of perfect matchings and Pauling's bond orders in hexagonal systems whose inner dual is a tree, in: *Graphs, Hypergraphs, and Applications*, Proceedings of the Conference on Graph Theory held in Eyba, October 1984, (Teubner-Texte zur Mathematik 73), BSB B. G. Teubner Verlagsgesellschaft Leipzig, Leipzig 1985, 80–91.

[7] P. John, and H. Sachs: Wegesysteme und Linearfaktoren in hexagonalen und quadratischen Systemen, in: *Graphen in Forschung und Unterricht* (Festschrift K. Wagner), Herausgeber R. Bodendiek et al., Verlag B. Franzbecker, Bad Salzdetfurth 1985, 85–101.

[8] P. John, H. Sachs, and H. Zernits: Counting perfect matchings in polyominoes, Proceedings of the International Conference "Combinatorial Analysis and Its Applications" (Pokrzywna, Poland, Sept. 16–20, 1985), *Zastosowania Matematyki* **19** (4), to appear.

[9] H. Sachs: Perfect matchings in hexagonal systems, *Combinatorica* **4** (1984), 89–99.

Kathie Cameron

Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario, Canada N2L 3G1

Horst Sachs

Technische Hochschule Ilmenau
Ilmenau, Germany