Zone Diagrams: Existence, Uniqueness and Algorithmic Challenge

Tetsuo Asano*
School of Information Science, JAIST
1-1 Asahidai, Nomi, Ishikawa
923-1292 Japan

Jiří Matoušek**
Department of Applied Mathematics and
Institute of Theoretical Computer Science (ITI)
Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic

Takeshi Tokuyama‡
Graduate School of Information Sciences
Tohoku University
Aramaki Aza Aoba, Aoba-ku, Sendai
980-8579 Japan

Abstract

A *zone diagram* is a new variation of the classical notion of Voronoi diagram. Given points (sites) $\mathbf{p}_1, \ldots, \mathbf{p}_n$ in the plane, each $\mathbf{p}_i$ is assigned a region R_i, but in contrast to the ordinary Voronoi diagrams, the union of the R_i has a nonempty complement, the *neutral zone*. The defining property is that each R_i consists of all $\mathbf{x} \in \mathbb{R}^2$ that lie closer (non-strictly) to $\mathbf{p}_i$ than to the union of all the other R_j, $j \neq i$. Thus, the zone diagram is defined implicitly, by a "fixed-point property," and neither its existence nor its uniqueness seem obvious. We establish existence using a general fixed-point result (a consequence of Schauder's theorem or Kakutani's theorem); this proof should generalize easily to related settings, say higher dimensions. Then we prove uniqueness of the zone diagram, as well as convergence of a natural iterative algorithm for computing it, by a geometric argument, which also relies on a result for the case of two sites in an earlier paper. Many challenging questions remain open.

*The part of this research by T.A. was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research on Priority Areas and Scientific Research (B).

**Parts of this research by J.M. were done during visits to the Japanese Advanced Institute for Science and Technology (JAIST) and to the ETH Zürich; the support of these institutions is gratefully acknowledged.

‡The part of this research by T.T. was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research on Priority Areas.

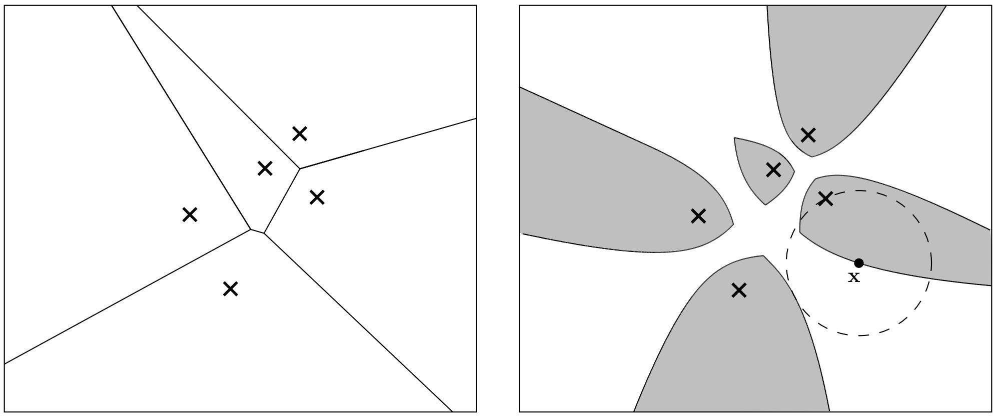

Figure 1: Five sites are marked by crosses. The left picture is the classical Voronoi diagram. The right picture shows the zone diagram: Each site $\mathbf{p}_i$ has a dominance region R_i, and the distance of each point $\mathbf{x}$ on the boundary of R_i to $\mathbf{p}_i$ equals the distance of $\mathbf{x}$ to the union of the other regions.

1 Introduction

Let us consider n points (sites) $\mathbf{p}_1, \ldots, \mathbf{p}_n$ in the plane. The left picture in Fig. 1 shows the (usual) Voronoi diagram, while the right picture is the *zone diagram*, a new notion investigated in the present paper.[1]

For a point $\mathbf{a}$ and a set $X \subseteq \mathbb{R}^2$ we define the *dominance region* of $\mathbf{a}$ with respect to X as

$$\operatorname{dom}(\mathbf{a}, X) = \{\mathbf{z} \in \mathbb{R}^2 : d(\mathbf{z}, \mathbf{a}) \leq d(\mathbf{z}, X)\},$$

where $d(\cdot, \cdot)$ denotes the Euclidean distance and $d(\mathbf{z}, X) = \inf_{\mathbf{x} \in X} d(\mathbf{z}, \mathbf{x})$. We note that $\operatorname{dom}(\mathbf{a}, X)$ is always convex and contains $\mathbf{a}$.

In the classical Voronoi diagram, the region of the site $\mathbf{p}_i$ is $\operatorname{dom}(\mathbf{p}_i, \{\mathbf{p}_j : j \neq i\})$, and the regions tile the whole plane. In a zone diagram, each $\mathbf{p}_i$ also has a region R_i, but the union of all the regions has a nonempty complement, called the *neutral zone*. We require that

$$R_i = \operatorname{dom}\Big(\mathbf{p}_i, \bigcup_{j \neq i} R_j\Big) \quad \text{for all } i = 1, 2, \ldots, n; \tag{1}$$

in words, the region of each site should consist of all points that are closer (non-strictly) to the site than to all of the other regions. In particular, each R_i is convex and contains $\mathbf{p}_i$.

The notion of zone diagram can be illustrated by a story on equilibrium in an "age of wars". There are n mutually hostile kingdoms. The ith kingdom has a castle at the site $\mathbf{p}_i$ and a territory R_i around it. The n territories are separated by a no-man's land, the neutral zone. If the territory R_i is attacked from another

[1] In earlier papers [2, 1] we have used the longer name *Voronoi diagram with neutral zone*, but here we propose the shorter term.

kingdom, an army departs from the castle $\mathbf{p}_i$ to intercept the attack. The interception succeeds if and only if the defending army arrives at the attacking point on the boundary of R_i sooner than the enemy. However, the attacker can secretly move his troops inside his territory, and the defense army can start from its castle only when the attacker leaves his territory. The zone diagram is an equilibrium configuration of the territories, such that every kingdom can guard its territory and no kingdom can grow without risk of invasion by other kingdoms.

The notion of zone diagram is, in our opinion, very interesting and it poses many mathematical and algorithmic challenges. Moreover, zone diagrams or variations could be useful for modeling natural phenomena. The classical Voronoi diagram, one of the basic geometric structures, appears in many fields and, among others, it is frequently used as a mathematical model of a simultaneous growth from several sites (cells in a tissue, a crystal lattice, geological patterns, regional equilibria in social sciences etc.). Voronoi diagrams and their numerous generalizations (see, e.g., [3, 5]) subdivide all of the space into dominance regions of the sites. However, geometric structures are sometimes observed in nature where the dominance regions do not cover everything, which might be a result of growth process where the growth terminates before the cell boundaries meet, due to some non-contact action.

The above definition of the zone diagram is implicit, since each region is defined in terms of the remaining ones. So it is not obvious whether any system of regions with the required property exists at all, and whether it is determined uniquely. Here we answer both of these questions affirmatively:

Theorem 1.1 *For every choice of n distinct sites $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^2$ there exists exactly one system $(R_1, \ldots, R_n)$ of subsets of the plane satisfying (1).*

Perhaps surprisingly, already the case of two sites ($n = 2$) is nontrivial. We showed existence and uniqueness for $n = 2$ in [1]. Here the two regions are mirror images of one another and they are bounded by an interesting curve called the *distance trisector curve*. We conjecture that this curve is not algebraic, and not even expressible by elementary functions (but we have no proof so far). On the other hand, its intersection with a given line parallel to the segment $\mathbf{p}_1\mathbf{p}_2$ can be computed to any desired precision in time polynomial in the number of required digits.

The *existence* part of Theorem 1.1 is proved in Section 4. We apply a well-known fixed-point theorem for infinite-dimensional Banach spaces to a suitable space of n-tuples of regions. This proof is conceptually simple and it appears quite robust, in the sense that it should be possible to adapt it to various natural generalizations of zone diagrams, such as zone diagrams in $\mathbb{R}^d$, zone diagrams of non-point sites, or α-zone diagrams (where each point of R_i should be α-times closer to $\mathbf{p}_i$ than to $\bigcup_{j \neq i} R_j$, for some real parameter $\alpha > 0$). We should remark, though, that such modifications are not necessarily trivial, since some elementary geometric estimates are needed that might prove technically challenging in some settings.

In Section 5 we prove the uniqueness in Theorem 1.1 and, at the same time, we also re-prove existence by a different method, similar to the one we used in

[1]. This method currently seems very much restricted to the planar case of zone diagrams, and several obstacles would have to be overcome before it could be generalized to $\mathbb{R}^3$, say.

In the uniqueness proof, we consider a natural iterative procedure for approximating the zone diagram. Let the sites $\mathbf{p}_1, \ldots, \mathbf{p}_n$ be fixed and let $\mathbf{R} = (R_1, \ldots, R_n)$ be an ordered n-tuple of regions (nonempty subsets of $\mathbb{R}^2$), where we assume $\mathbf{p}_1 \in R_1, \ldots, \mathbf{p}_n \in R_n$. We define $\mathbf{Dom}(\mathbf{R})$ as the ordered n-tuple $\mathbf{S} = (S_1, S_2, \ldots, S_n)$ of new regions, where $S_i = \mathrm{dom}(\mathbf{p}_i, \bigcup_{j \neq i} R_j)$. Thus, rephrasing our definition of a zone diagram, an n-tuple $\mathbf{R} = (R_1, \ldots, R_n)$ of regions is a zone diagram of $\mathbf{p}_1, \ldots, \mathbf{p}_n$ if it is a fixed point of the operator $\mathbf{Dom}$; that is, if $\mathbf{R} = \mathbf{Dom}(\mathbf{R})$.

For two n-tuples $\mathbf{R} = (R_1, \ldots, R_n)$ and $\mathbf{S} = (S_1, \ldots, S_n)$, let us write $\mathbf{R} \preceq \mathbf{S}$ if $R_i \subseteq S_i$ for all $i = 1, 2, \ldots, n$. It is immediate from the definition that if $\mathbf{R} \preceq \mathbf{S}$, then $\mathbf{Dom}(\mathbf{R}) \succeq \mathbf{Dom}(\mathbf{S})$ (that is, the dominance operator is antimonotone with respect to $\preceq$). Let $\mathbf{I}^{(0)} = (\{\mathbf{p}_1\}, \ldots, \{\mathbf{p}_n\})$ be the (smallest possible) system of one-point regions, let $\mathbf{O}^{(0)} = (O_1^{(0)}, \ldots, O_n^{(0)}) = \mathbf{Dom}(\mathbf{I}^{(0)})$ be the regions of the classical Voronoi diagram of $\mathbf{p}_1, \ldots, \mathbf{p}_n$, and for $k = 1, 2, \ldots$ we inductively define $\mathbf{I}^{(k)} = \mathbf{Dom}(\mathbf{O}^{(k-1)})$, $\mathbf{O}^{(k)} = \mathbf{Dom}(\mathbf{I}^{(k)})$.

Antimonotonicity of $\mathbf{Dom}$ and induction yield $\mathbf{I}^{(0)} \preceq \mathbf{I}^{(1)} \preceq \mathbf{I}^{(2)} \preceq \cdots$ and $\mathbf{O}^{(0)} \succeq \mathbf{O}^{(1)} \succeq \mathbf{O}^{(2)} \succeq \cdots$. Moreover, if $\mathbf{R}$ is a zone diagram, i.e., satisfies $\mathbf{R} = \mathbf{Dom}(\mathbf{R})$, then we have $\mathbf{I}^{(0)} \preceq \mathbf{R}$ by definition, and induction and antimonotonicity give $\mathbf{I}^{(k)} \preceq \mathbf{R} \preceq \mathbf{O}^{(k)}$ for all k. The $\mathbf{I}^{(k)}$ form an increasing sequence of inner approximations of the zone diagram, while the $\mathbf{O}^{(k)}$ form a decreasing sequence of outer approximations; see Fig. 2.

In Section 5 we show that the inner and outer approximations converge to the same limit, which has to be the unique zone diagram. This also gives a quite practical algorithm for approximate construction of the zone diagram. The regions of the $\mathbf{I}^{(k)}$ and $\mathbf{O}^{(k)}$ can be approximated by convex polygons with many sides—this is how the pictures of zone diagrams in this paper were obtained. With some care in implementation one can actually get pairs of polygons that are provably inner and outer approximations, respectively, of the regions of the zone diagram. Experiments indicate that the convergence of this algorithm is quite fast, at least for small sets of sites (each iteration is computationally demanding, though). Unfortunately, we have no theoretical estimate of the convergence rate of this algorithm. An example illustrating some of the difficulties in proving estimates is given in Section 6. We also mention some additional results and questions there.

2 A Guided Tour of Zone Diagrams

Before we start with proofs, we explain, mainly by pictures, some interesting phenomena arising in zone diagrams, illustrating that they behave very differently from the classical Voronoi diagrams.

The left picture in Fig. 3 shows the zone diagram of two sites (the distance trisector curve), and the right picture shows the zone diagram after adding a third site marked by a small disk. The boundary curves of the regions from the

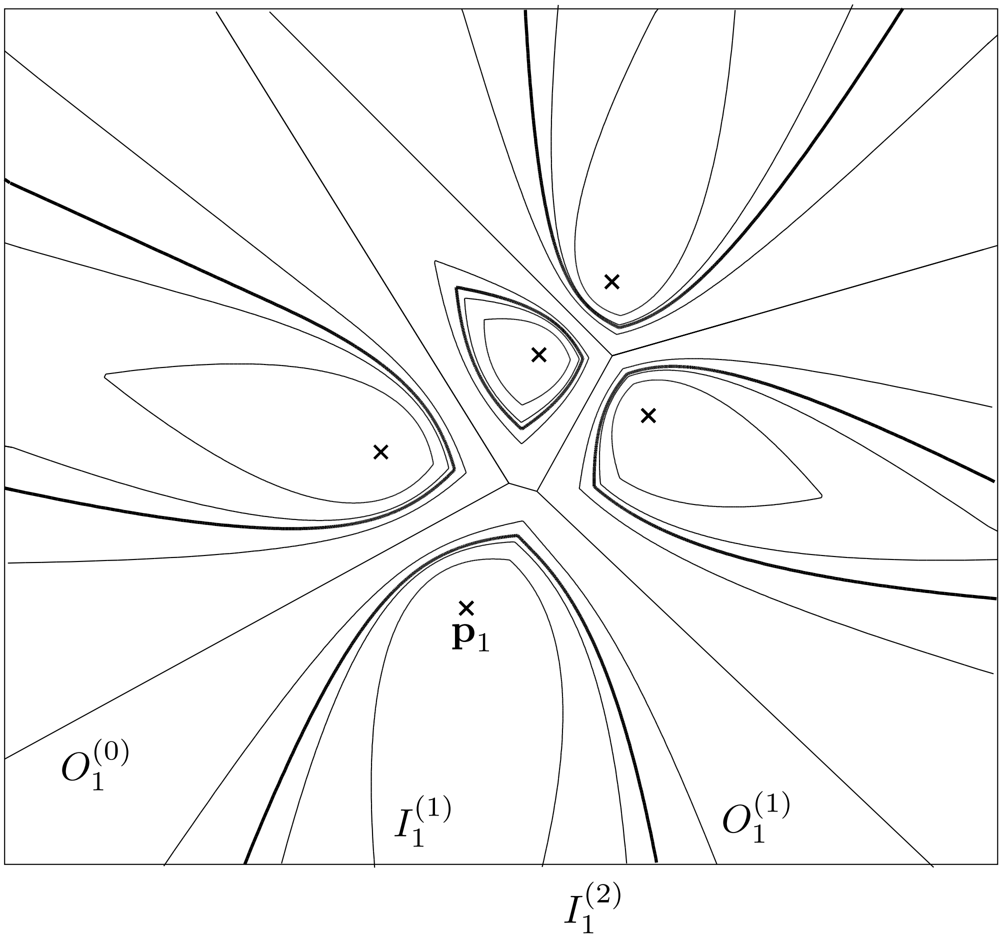

Figure 2: The inner and outer approximations $\mathbf{I}^{(k)}$ and $\mathbf{O}^{(k)}$. In particular, $\mathbf{O}^{(0)}$ forms the classical Voronoi diagram.

previous 2-site diagram are also shown, and one can see that the region of the top site has *gained* new area after the new site was added (this cannot happen in classical Voronoi diagrams). This is very intuitive in the war interpretation: The animosity of the two nearby sites weakens them and the top site gets relatively stronger.

In a classical Voronoi diagram for sites $\mathbf{p}_1, \ldots, \mathbf{p}_n$, the region of $\mathbf{p}_i$ is the intersection of the regions of $\mathbf{p}_i$ in the two-site Voronoi diagrams for all pairs $\{\mathbf{p}_i, \mathbf{p}_j\}$, $j \neq i$. Consequently, each region is bounded by segments that arise as bisectors of pairs of sites. Fig. 3 illustrates that no analogy holds for zone diagrams. Indeed, segments of the distance trisector curve do appear as portions of the boundary of the regions for 3 sites, but we can also have other kinds of curves.

In Fig. 3 we can see a simple instance of this (straight segments appear near the bottom tip of the top region); other examples exhibit more complicated curves as well. The proof in Section 5 tells something about the nature of all

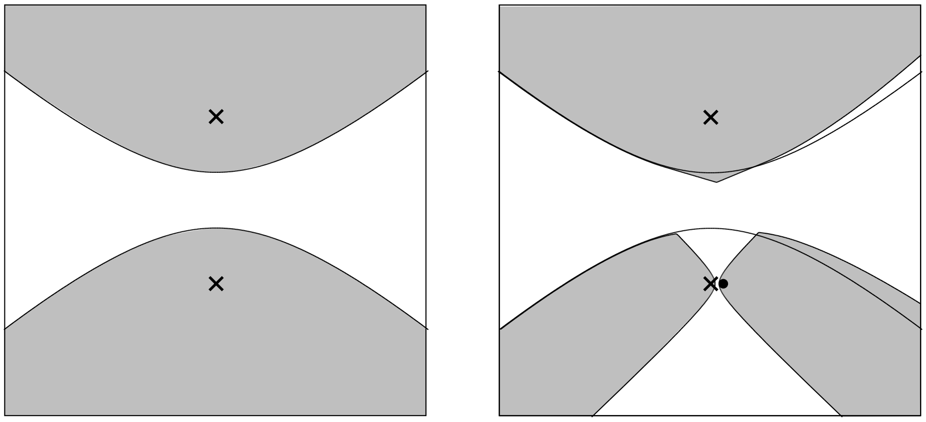

Figure 3: The zone diagram of two sites (left) and the zone diagram after adding a third site marked by • (right). The top site gains area, and the regions are not bounded only by arcs of distance trisector curves.

curves that can ever appear, but some interesting questions remain open.

The left picture in Fig. 4 shows an aesthetically pleasing zone diagram. In this case all of the regions are bounded, which again doesn't happen in classical Voronoi diagrams. Such "flowers" scaled down to a tiny size can be used in constructing examples; the right picture shows a small flower and an isolated site $\mathbf{q}$. As the flower gets smaller, the region of $\mathbf{q}$ approaches a halfplane, that is, the region of $\mathbf{q}$ in a two-site classical Voronoi diagram with a single site at the center of the flower.

3 Preliminaries

Here we introduce some notation and some simple and/or known facts.

We note that for any $X \subseteq \mathbb{R}^2$ the dominance region $\mathrm{dom}(\mathbf{a}, X)$ is a closed convex set, since it can be represented as the intersection $\bigcap_{\mathbf{x}\in X} \mathrm{dom}(\mathbf{a}, \{\mathbf{x}\})$ of halfplanes.

The boundary of a set $X \subseteq \mathbb{R}^2$ is denoted by ∂X.

In analogy to the dominance region notation $\mathrm{dom}(\mathbf{a}, X)$ we will also use the *bisector* notation defined by $\mathrm{bisect}(\mathbf{a}, X) = \{\mathbf{z} \in \mathbb{R}^2 : d(\mathbf{z}, \mathbf{a}) = d(\mathbf{z}, X)\}$.

For a nonempty closed convex set $C \subseteq \mathbb{R}^2$ and a point $\mathbf{x} \in \mathbb{R}^2$ we let $\mathrm{prox}_C(\mathbf{x})$ denote the point of C nearest to $\mathbf{x}$. It is well known that this point is unique. Moreover, for C fixed, the mapping prox_C (the *metric projection*) is continuous, and actually 1-Lipschitz.

We will need the following lemma, expressing a kind of continuity of the dominance operator.

Lemma 3.1 *Let* $\mathbf{a} \in \mathbb{R}^2$ *be a point and let* $X_1 \supseteq X_2 \supseteq X_3 \supseteq \cdots$ *be a decreasing*

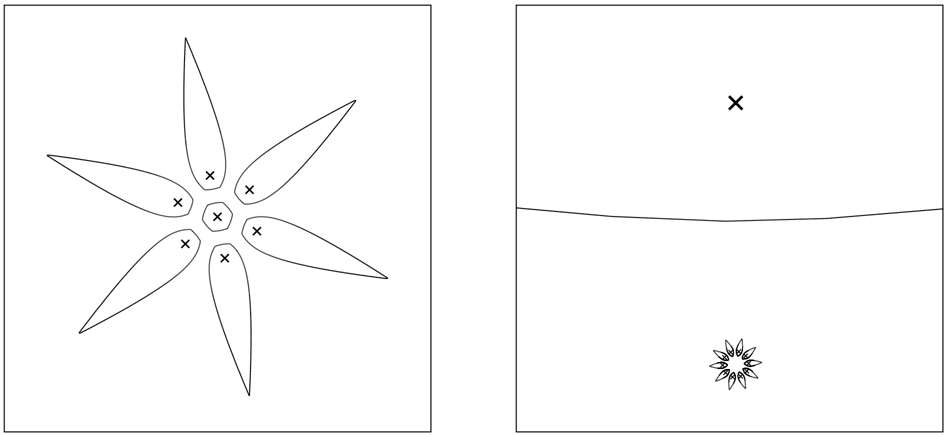

Figure 4: A flower (left); a small flower induces almost the classical Voronoi region of an isolated site (right).

sequence of closed subsets of $\mathbb{R}^2$ *with* $\mathbf{a} \notin X_1$. *Let us set* $X = \bigcap_{k=1}^{\infty} X_k$. *Then*

$$\mathrm{dom}(\mathbf{a}, X) = \mathrm{cl}\Big(\bigcup_{k=1}^{\infty} \mathrm{dom}(\mathbf{a}, X_k)\Big),$$

where cl(.) *denotes the topological closure.*

Proof. The inclusion "$\supseteq$" is clear from $X \subseteq X_k$ for all k and antimonotonicity of dom(.). To prove the opposite inclusion, we fix $\mathbf{x} \in \mathrm{dom}(\mathbf{a}, X)$ arbitrarily, we choose $\varepsilon > 0$ arbitrarily small, and we show that there exists $k = k(\mathbf{x}, \varepsilon)$ with $d(\mathbf{x}, \mathrm{dom}(\mathbf{a}, X_k)) < \varepsilon$. We may assume $\mathbf{x} \neq \mathbf{a}$, for otherwise, we even have $\mathbf{x} \in \mathrm{dom}(\mathbf{a}, X_1)$.

Since $\mathbf{a} \notin X_1$ and X_1 is closed, we have $\delta = d(\mathbf{a}, X_1) > 0$. The set X lies outside the region shown in Fig. 5. Elementary geometric considerations show that all interior points $\mathbf{y}$ of the segment $\mathbf{ax}$ satisfy $d(\mathbf{y}, \mathbf{a}) < d(\mathbf{y}, X)$. Let us choose such a point $\mathbf{y}$ with $d(\mathbf{y}, \mathbf{x}) < \varepsilon$.

A simple compactness argument, which we omit, shows that for any point $\mathbf{q}$ we have $d(\mathbf{q}, X) = \lim_{k\to\infty} d(\mathbf{q}, X_k)$. Hence there exists k with $d(\mathbf{y}, \mathbf{a}) < d(\mathbf{y}, X_k)$, and thus $\mathbf{y} \in \mathrm{dom}(\mathbf{a}, X_k)$. Hence $d(\mathbf{x}, \mathrm{dom}(\mathbf{a}, X_k)) < \varepsilon$ as claimed.

□

4 Existence of the Zone Diagram

In this section we prove the existence of (at least one) zone diagram for every set $\{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ of distinct sites in the plane. Let $\mathcal{R}$ denote the set of all n-tuples $\mathbf{R} = (R_1, \ldots, R_n)$ of sets with $\mathbf{p}_i \in R_i \subseteq \mathbb{R}^2$.

Plan of the proof. We want to show the existence of a fixed point of the dominance operator $\mathbf{Dom}\colon \mathcal{R} \to \mathcal{R}$ (defined in Section 1). We are going to apply the following theorem (which can be seen as a special case of two famous

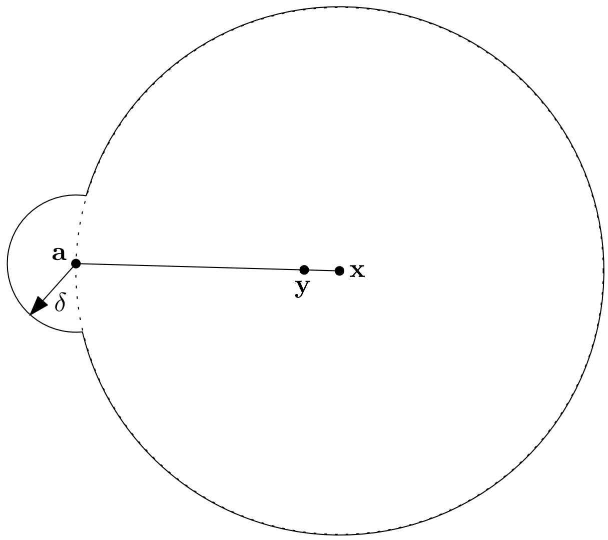

Figure 5: Illustration to the proof of Lemma 3.1.

theorems in fixed-point theory, Schauder's and Kakutani's; see, for example, Zeidler [6], Corollary 2.13):

Theorem 4.1 *Let Z be a Banach space, let $K \subset Z$ be a nonempty, compact, and convex set, and let $F: K \to K$ be a continuous map. Then F has at least one fixed point.*

In our application of this fixed-point theorem, we will define a suitable set $\mathcal{S} \subseteq \mathcal{R}$ of n-tuples of regions, and we will define an embedding $\varphi: \mathcal{S} \to Z$ for a suitable Banach space Z. The image $\varphi(\mathcal{S})$ will play the role of K in the fixed-point theorem, and F is the mapping $K \to Z$ corresponding to **Dom** under φ (formally, $F = \varphi \circ \mathbf{Dom} \circ \varphi^{-1}$). We thus need to verify that K is convex and compact, that $F(K) \subseteq K$, and that F is continuous.

Here we will present our original "manual" approach to this task. We will define $\mathcal{S}$ in a slightly tricky manner, which makes the verification of the above conditions quite easy, *except* for checking the continuity of F, which is not really hard but we need about 2 pages of elementary geometric arguments and estimates.

An alternative strategy. An alternative, somewhat simpler and, in a sense, more natural approach (leading to a formally slightly weaker result) was suggested to us by Eva Kopecká. We sketch it here and then we return to our original proof. First of all, we restrict everything to a bounded region Q, say a large square containing all the sites, and prove the existence of the zone diagram only in this region (that's why the result is formally weaker). Then we let Q_i be the intersection of Q with the cell of $\mathbf{p}_i$ in the classical Voronoi diagram of $\mathbf{p}_1, \ldots, \mathbf{p}_n$, and we define $\mathcal{S}$ as the set of all n-tuples $(S_1, \ldots, S_n)$ of nonempty closed sets with $\mathbf{p}_i \in S_i \subseteq Q_i$. We equip this $\mathcal{S}$ with the Hausdorff distance metric; formally, the distance of $(S_1, \ldots, S_n)$ and $(S'_1, \ldots, S'_n)$

equals $\max_{i=1,2,\ldots,n} h(S_i, S_i')$, where h is the Hausdorff distance. It follows from the work of Curtis, Schori and West from the 1970s (culminating in [4], where other references can also be found) that this $\mathcal{S}$ as a topological space is homeomorphic to the Hilbert cube, which is a compact convex subset of ℓ_2. Hence for application of Theorem 4.1 it is enough to verify that **Dom** maps $\mathcal{S}$ into $\mathcal{S}$ (clear) and that it is continuous with respect to the Hausdorff metric. This is similar in spirit to our continuity argument below but simpler.

Radial functions. We return to our original approach. Let S^1 denote the unit circle; we will interpret its points as angles in the interval $[0, 2\pi)$. We will call a continuous function $\rho\colon S^1 \to [0, \frac{\pi}{2}]$ a *radial function.* For such a ρ and a point $\mathbf{p} \in \mathbb{R}^2$, we define a star-shaped region $R = \mathrm{reg}_{\mathbf{p}}(\rho)$ such that the ray emanating from $\mathbf{p}$ at angle α intersects R in a segment of length $\tan(\rho(\alpha))$. Formally,

$$\mathrm{reg}_{\mathbf{p}}(\rho) = \bigcup_{\alpha\in[0,2\pi)} \mathbf{p}\mathbf{x}_\alpha, \quad \text{where } \mathbf{x}_\alpha = \mathbf{p} + \tan(\rho(\alpha))(\cos\alpha, \sin\alpha)$$

(if $\rho(\alpha) = \frac{\pi}{2}$, then $\mathbf{p}\mathbf{x}_\alpha$ is defined as the full semiinfinite ray). We note that the length of the segment in direction α is not $\rho(\alpha)$ but rather $\tan(\rho(\alpha))$. This ensures that we deal with bounded radial functions, although the considered planar regions are often unbounded. The choice of the tangent function to map a bounded interval to $[0, \infty)$ is somewhat arbitrary, but certainly not every function would do. For example, we have to be careful about how we measure the distance of regions, in order to obtain continuity of the operator **Dom**.

For simplicity, let us assume that every two sites $\mathbf{p}_i \neq \mathbf{p}_j$ have distance at least 4 (this will save us one parameter, standing for the minimum distance of sites, in the forthcoming calculations).

Now we can define our Banach space and the set K.

Definition 4.2 *Let Z denote the Banach space of all n-tuples $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)$ of continuous functions $\rho_i\colon S^1 \to \mathbb{R}$, endowed with the supremum norm: $\|\boldsymbol{\rho}\|_\infty = \max_{i=1,2,\ldots,n} \max_{\alpha\in S^1} |\rho_i(\alpha)|$.*

Let $K \subset Z$ consist of all $\boldsymbol{\rho} \in Z$ satisfying the following conditions:

(i) *The image of each ρ_i is contained in $[0, \frac{\pi}{2}]$ (that is, ρ_i is a radial function).*

(ii) *We have $\mathbf{I}^{(1)} \preceq \mathbf{reg}(\boldsymbol{\rho}) \preceq \mathbf{O}^{(1)}$, where $\mathbf{reg}(\boldsymbol{\rho}) = (\mathrm{reg}_{\mathbf{p}_1}(\rho_1), \ldots, \mathrm{reg}_{\mathbf{p}_n}(\rho_n)) \in \mathcal{R}$ is the system of regions defined by $\boldsymbol{\rho}$ (the componentwise inclusion operator $\preceq$ and the $\mathbf{I}^{(k)}$ and $\mathbf{O}^{(k)}$ were introduced in Section 1). (We note that this simply means pointwise lower and upper bounds on each ρ_i.)*

(iii) *Each ρ_i is 2-Lipschitz.*

Further, we set $\mathcal{S} = \mathbf{reg}(K)$ (so $\mathbf{reg}$ plays the role of φ^{-1} in the abstract outline of the argument given above).

The set K is clearly nonempty and convex (since convex combinations preserve the conditions in the definition of K), and it is easily seen to be compact by the Arzèla–Ascoli theorem, which implies, in particular, that any closed set of uniformly bounded 2-Lipschitz functions on a compact set is compact.

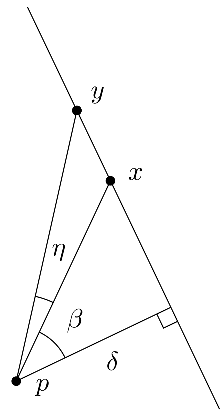

Figure 6: The radial function of a convex set containing the unit disk is Lipschitz.

Lemma 4.3 *For every n-tuple $\mathbf{R} \in \mathcal{S}$ of regions we have $\mathbf{Dom}(\mathbf{R}) \in \mathcal{S}$. Consequently, the mapping $F\colon K \to K$ given by $F = \mathbf{reg}^{-1} \circ \mathbf{Dom} \circ \mathbf{reg}$ is well defined.*

Proof. Let $\mathbf{S} = \mathbf{Dom}(\mathbf{R}) = (S_1, \ldots, S_n)$. Each S_i is convex and hence given by a radial function, so $\boldsymbol{\sigma} = \mathbf{reg}^{-1}(\mathbf{S})$ is well defined.

Since $\mathbf{R} \in \mathcal{S}$, we have $\mathbf{I}^{(1)} \preceq \mathbf{R} \preceq \mathbf{O}^{(1)}$, and by antimonotonicity we get $\mathbf{I}^{(2)} = \mathbf{Dom}(\mathbf{O}^{(1)}) \preceq \mathbf{Dom}(\mathbf{R}) = \mathbf{S}$. Similarly $\mathbf{S} \preceq \mathbf{O}^{(2)}$, so $\mathbf{I}^{(1)} \preceq \mathbf{I}^{(2)} \preceq \mathbf{S} \preceq \mathbf{O}^{(2)} \preceq \mathbf{O}^{(1)}$, and $\mathbf{S}$ satisfies (ii) from Definition 4.2.

For proving (iii), we first note that, assuming $d(\mathbf{p}_i, \mathbf{p}_j) \geq 4$ for all $i \neq j$, each $I_i^{(1)}$ contains the unit disk centered at $\mathbf{p}_i$. Indeed, the regions of $\mathbf{O}^{(0)}$ are the classical Voronoi regions, and so $I_i^{(1)}$ consists of the points closer to $\mathbf{p}_i$ than to the Voronoi regions of the other points. It remains to note that the Voronoi region of $\mathbf{p}_i$ contains the disk of radius 2 around $\mathbf{p}_i$.

It remains to prove the following claim: *If ρ is a radial function, $\mathbf{p} \in \mathbb{R}^2$, and the set $R = \mathrm{reg}_{\mathbf{p}}(\rho)$ is convex and contains the unit disk centered at $\mathbf{p}$, then ρ is a 2-Lipschitz function.* (Actually, an easy refinement of the proof below yields 1-Lipschitz.)

First we note that R may be assumed to be bounded, since intersecting R with a very large disc changes the radial function by an arbitrarily small amount.

Let $\mathbf{x}$ be the boundary point of R in direction α and let $\mathbf{y}$ be the boundary point in direction $\alpha + \eta$, where (as we may assume) $0 < \eta \leq \frac{\pi}{6}$, say (see Fig. 6). Then it is easy to see that the line $\mathbf{xy}$ has distance $\delta \geq \frac{1}{2}$ to $\mathbf{p}$: We have $|\mathbf{px}| = \delta/\cos\beta$ and $|\mathbf{py}| = \delta/\cos(\beta+\eta)$, and thus $|\rho(\alpha+\eta) - \rho(\alpha)| = |\arctan(\delta/\cos(\beta+\eta)) - \arctan(\delta/\cos\beta)| \leq \eta/\delta$, where the last inequality follows from the Mean Value Theorem applied to the function $\beta \mapsto \arctan(\delta/\cos\beta)$. Its first derivative is $\frac{1}{\delta}\sin\beta/(1 + (\cos\beta)^2/\delta^2)$ and this is obviously bounded by $\frac{1}{\delta}$. $\square$

In order to apply Theorem 4.1, it thus remains to prove the following.

Lemma 4.4 *The mapping $F = \mathbf{reg}^{-1} \circ \mathbf{Dom} \circ \mathbf{reg} \colon K \to K$ is continuous.*

We give the proof of the above lemma in the appendix. The existence of a zone diagram then follows from Theorem 4.1.

5 Uniqueness of the zone diagram

In this section we prove both existence and uniqueness of the zone diagram for any n distinct sites $\mathbf{p}_1, \ldots, \mathbf{p}_n$, as well as convergence of the iterative procedure described in the introduction. The proof is divided into two steps. The first step is the following quite intuitive statement:

Lemma 5.1 *Let $\mathbf{I}^{(k)} = (I_1^{(k)}, \ldots, I_n^{(k)})$ be the inner approximations of the zone diagram and let $\mathbf{O}^{(k)} = (O_1^{(k)}, \ldots, O_n^{(k)})$ be the outer approximations. For $i = 1, 2, \ldots, n$ let us set*

$$I_i = \mathrm{cl}\Big(\bigcup_{k=0}^{\infty} I_i^{(k)}\Big), \quad O_i = \bigcap_{k=0}^{\infty} O_i^{(k)}.$$

Then $\mathbf{I} = (I_1, \ldots, I_n)$ and $\mathbf{O} = (O_1, \ldots, O_n)$ satisfy $\mathbf{I} = \mathbf{Dom}(\mathbf{O})$ and $\mathbf{O} = \mathbf{Dom}(\mathbf{I})$.

Proof. This statement is not as obvious as it might perhaps seem. First we check $\mathbf{O} = \mathbf{Dom}(\mathbf{I})$; this is entirely straightforward. Fixing i, we want to verify

$$O_i = \mathrm{dom}\Big(\mathbf{p}_i, \bigcup_{j \neq i} I_j\Big). \tag{2}$$

Since for every k, $I_j \supseteq I_j^{(k)}$, we have $\mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} I_j) \subseteq \mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} I_j^{(k)}) = O_i^{(k)}$, and thus $\mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} I_j) \subseteq \bigcap_{k=0}^{\infty} O_i^{(k)} = O_i$; this is "$\supseteq$" in (2). For the converse inclusion, we assume $\mathbf{x} \notin \mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} I_j)$. Then there exists j_0 and $\mathbf{y} \in I_{j_0}$ with $d(\mathbf{x}, \mathbf{p}_i) > d(\mathbf{x}, \mathbf{y})$. Setting $\varepsilon = d(\mathbf{x}, \mathbf{p}_i) - d(\mathbf{x}, \mathbf{y})$, we can choose k sufficiently large so that $I_{j_0}^{(k)}$ contains a point $\mathbf{y}'$ with $d(\mathbf{y}, \mathbf{y}') < \varepsilon$, and then $d(\mathbf{x}, \mathbf{y}') \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{y}') < d(\mathbf{x}, \mathbf{y}) + \varepsilon = d(\mathbf{x}, \mathbf{p}_i)$. Hence $\mathbf{x} \notin \mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} I_j^{(k)}) = O_i^{(k)}$, and $\mathbf{x} \notin O_i$ either. This proves (2).

We now turn to showing $\mathbf{I} = \mathbf{Dom}(\mathbf{O})$; that is,

$$I_i = \mathrm{dom}\Big(\mathbf{p}_i, \bigcup_{j \neq i} O_j\Big).$$

Here "$\subseteq$" is again straightforward, but "$\supseteq$" needs more properties of dom(.). We apply Lemma 3.1 with $\mathbf{a} = \mathbf{p}_i$ and $X_k = \bigcup_{j\neq i} O_j^{(k)}$. Since the $O_i^{(1)}$ are disjoint, we have $X = \bigcap_{k=1}^{\infty} X_k = \bigcup_{j\neq i} \bigcap_{k=1}^{\infty} O_j^{(k)} = \bigcup_{j\neq i} O_j$, and so the lemma tells us that $\mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} O_j) = \mathrm{cl}(\bigcup_{k=1}^{\infty} \mathrm{dom}(\mathbf{p}_i, \bigcup_{j\neq i} O_j^{(k)})) = \mathrm{cl}(\bigcup_{k=1}^{\infty} I_i^{(k+1)}) = I_i$ as required. Lemma 5.1 is proved. $\square$

In the second step, which is the essence of the proof, we establish the following:

Proposition 5.2 *Let the sites* $\mathbf{p}_1, \ldots, \mathbf{p}_n$ *be fixed and let* $\mathbf{S} = (S_1, \ldots, S_n)$ *and* $\mathbf{T} = (T_1, \ldots, T_n)$ *be* n*-tuples of regions satisfying* $\mathbf{S} = \mathbf{Dom}(\mathbf{T})$ *and* $\mathbf{T} = \mathbf{Dom}(\mathbf{S})$. *Then* $\mathbf{S} = \mathbf{T}$, *and consequently,* $\mathbf{S}$ *is a zone diagram of* $\mathbf{p}_1, \ldots, \mathbf{p}_n$.

Proof of Theorem 1.1. Let $\mathbf{I}$ and $\mathbf{O}$ be as in Lemma 5.1. Then Proposition 5.2 with $\mathbf{S} = \mathbf{I}$ and $\mathbf{T} = \mathbf{O}$ shows that $\mathbf{I} = \mathbf{O}$ is a zone diagram. Moreover, if $\mathbf{R}$ is any zone diagram of $\mathbf{p}_1, \ldots, \mathbf{p}_n$, we have $\mathbf{I}^{(k)} \preceq \mathbf{R} \preceq \mathbf{O}^{(k)}$ for all k as was explained in Section 1. Hence $\mathbf{I} \preceq \mathbf{R} \preceq \mathbf{O}$ and $\mathbf{R} = \mathbf{I} = \mathbf{O}$. Thus the zone diagram is unique. □

Preparations for the proof of Proposition 5.2. We assume that $\mathbf{S}$ and $\mathbf{T}$ with $\mathbf{S} = \mathbf{Dom}(\mathbf{T})$ and $\mathbf{T} = \mathbf{Dom}(\mathbf{S})$ have been fixed. Since each S_i and T_i is a dominance region, it is a closed convex set. Since $\mathbf{I}^{(1)} \preceq \mathbf{S}, \mathbf{T}$, each S_i and T_i contains a small open disk around $\mathbf{p}_i$. Moreover, each S_i is disjoint from all T_j, $j \neq i$, and vice versa.

We introduce the following terminology: Let $\mathbf{a} \in S_1$ be a point. The *nearest points* of $\mathbf{a}$ are the points of $\bigcup_{i=2}^n T_i$ with the minimum distance to $\mathbf{a}$. Since each T_i is convex, it contains at most one of the nearest points of $\mathbf{a}$. The point $\mathbf{a}$ is called a *singular point* if it has more than one nearest point; otherwise, it is called a *regular point*. All of this refers to the situation $\mathbf{a} \in S_1$; if we speak about nearest points of some $\mathbf{a} \in T_2$, say, we mean the points of $\bigcup_{i \neq 2} S_i$ with the smallest distance to $\mathbf{a}$.

Let $\breve{\mathbf{a}} \in \bigcup_{i=2}^n T_i$ be a nearest point of $\mathbf{a} \in \partial S_1$. We call it *visible* if the segment $\mathbf{a}\breve{\mathbf{a}}$ intersects S_1 only at $\mathbf{a}$, and we call it *obscured* otherwise.

Lemma 5.3

(i) *Let* $\mathbf{a} \in \partial S_1$ *and let* $\breve{\mathbf{a}} \in \bigcup_{i=2}^n T_i$ *be a nearest point of it, say with* $\breve{\mathbf{a}} \in T_2$. *Then* $d(\mathbf{p}_1, \mathbf{a}) = d(\mathbf{a}, \breve{\mathbf{a}}) \geq d(\breve{\mathbf{a}}, \mathbf{p}_2)$.

(ii) *In the setting of (i),* $\breve{\mathbf{a}}$ *is visible.*

(iii) *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be distinct boundary points of* S_1, *let* $\breve{\mathbf{a}}$ *be a nearest point of* $\mathbf{a}$ *and* $\breve{\mathbf{b}}$ *a nearest point of* $\mathbf{b}$. *Then the segments* $\mathbf{a}\breve{\mathbf{a}}$ *and* $\mathbf{b}\breve{\mathbf{b}}$ *do not intersect, except possibly if* $\breve{\mathbf{a}} = \breve{\mathbf{b}}$.

Proof. In part (i) we have $d(\mathbf{p}_1, \mathbf{a}) \leq d(\mathbf{a}, \breve{\mathbf{a}})$ because $\mathbf{a} \in S_1 = \mathrm{dom}(\mathbf{p}_1, \bigcup_{i=2}^n T_i) \subseteq \mathrm{dom}(\mathbf{p}_1, \{\breve{\mathbf{a}}\})$. If we had $d(\mathbf{p}_1, \mathbf{a}) < d(\mathbf{a}, \breve{\mathbf{a}})$, then a small neighborhood of $\mathbf{a}$ would be contained in S_1 as well, but $\mathbf{a}$ is a boundary point. Next, we have $\breve{\mathbf{a}} \in T_2 = \mathrm{dom}(\mathbf{p}_2, \bigcup_{i=2}^n S_i) \subseteq \mathrm{dom}(\mathbf{p}_2, \{\mathbf{a}\})$, and $d(\mathbf{a}, \breve{\mathbf{a}}) \geq d(\mathbf{p}_2, \breve{\mathbf{a}})$ follows, which finishes the proof of (i).

For (ii), let us suppose that $\mathbf{a} \in \partial S_1$ has a nearest point $\breve{\mathbf{a}} \in T_i$ and that $\mathbf{b} \neq \mathbf{a}$ is a point of S_1 on the segment $\mathbf{a}\breve{\mathbf{a}}$. Then $d(\breve{\mathbf{a}}, \mathbf{b}) = d(\breve{\mathbf{a}}, \mathbf{a}) - d(\mathbf{a}, \mathbf{b}) = d(\mathbf{p}_1, \mathbf{a}) - d(\mathbf{a}, \mathbf{b})$. On the other hand, the point $\mathbf{b}$ doesn't lie on the segment $\mathbf{p}_1\mathbf{a}$ (otherwise, $\mathbf{p}_1$ would be closer to $\mathbf{a}$ than $\breve{\mathbf{a}}$), and hence we have the strict triangle inequality $d(\mathbf{b}, \mathbf{p}_1) > d(\mathbf{a}, \mathbf{p}_1) - d(\mathbf{a}, \mathbf{b})$. Thus $d(\breve{\mathbf{a}}, \mathbf{b}) < d(\mathbf{b}, \mathbf{p}_1)$, contradicting to $\mathbf{b} \in S_1$.

In (iii), we suppose for contradiction that $\breve{\mathbf{a}} \neq \breve{\mathbf{b}}$ and the segments $\mathbf{a}\breve{\mathbf{a}}$ and $\mathbf{b}\breve{\mathbf{b}}$ cross at a point $\mathbf{x}$. Then $d(\mathbf{a}, \breve{\mathbf{b}}) \leq d(\mathbf{a}, \mathbf{x}) + d(\mathbf{x}, \breve{\mathbf{b}})$ and $d(\mathbf{b}, \breve{\mathbf{a}}) \leq$

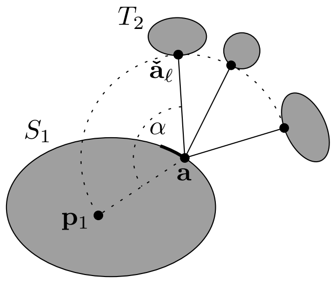

Figure 7: The left nearest point of $\mathbf{a}$.

$d(\mathbf{b}, \mathbf{x}) + d(\mathbf{x}, \check{\mathbf{a}})$ by the triangle inequality, with at least one of the inequalities strict. The two right-hand sides together equal $d(\mathbf{a}, \check{\mathbf{a}}) + d(\mathbf{b}, \check{\mathbf{b}})$. Then $\check{\mathbf{a}}$ cannot be a nearest point of $\mathbf{a}$ or $\check{\mathbf{b}}$ cannot be a nearest point of $\mathbf{b}$. The lemma is proved. $\square$

Let $\mathbf{a}$ be a boundary point of S_1. For each of the nearest points $\check{\mathbf{a}}$ of $\mathbf{a}$, we consider the angle $\alpha = \angle \check{\mathbf{a}}\mathbf{a}\mathbf{p}_1$ (measured counterclockwise; $0 < \alpha < 2\pi$). The *left nearest point* of $\mathbf{a}$ is the one with α minimum. We denote it by $\check{\mathbf{a}}_\ell$; see Fig. 7.

We make the following convention: Let $\mathbf{a}$ and $\mathbf{b}$ be two points on the boundary of some S_i or T_i. We say that $\mathbf{b}$ lies *left* of $\mathbf{a}$ if the angle $\angle \mathbf{a}\mathbf{p}_i\mathbf{b}$, measured counterclockwise, is between 0 and π (this will always concern very close points $\mathbf{a}$ and $\mathbf{b}$, and then we see $\mathbf{b}$ on the left of $\mathbf{a}$ when looking from $\mathbf{p}_i$).

Lemma 5.4 *Let $\mathbf{a} \in \partial S_1$, let $\check{\mathbf{a}}_\ell$ be the left nearest point of $\mathbf{a}$, and assume $\check{\mathbf{a}}_\ell \in T_2$. Then there exists a neighborhood U of $\mathbf{a}$ such that all points of ∂S_1 lying left of $\mathbf{a}$ and in U have exactly one nearest point, and moreover, this nearest point lies on ∂T_2, on the right of $\check{\mathbf{a}}_\ell$, and near to it (as near as desired if U is chosen small enough).*

Proof. First we note that if there are points $\mathbf{x} \in \partial S_1$ arbitrarily near to $\mathbf{a}$ that have a nearest point in some T_j, then by the continuity of the metric projection, $\mathbf{a}$ has a nearest point in T_j as well. Thus, supposing for contradiction that an arbitrarily small left neighborhood of $\mathbf{a}$ in ∂S_1 contained a point $\mathbf{x}$ with a nearest point $\check{\mathbf{x}}$ in T_j, $j \neq 2$, then $\check{\mathbf{x}}$ would get arbitrarily close to the nearest point of $\mathbf{a}$ in T_j. But then, if the neighborhood is taken sufficiently small, the segment $\mathbf{x}\check{\mathbf{x}}$ has to cross the segment $\mathbf{a}\check{\mathbf{a}}_\ell$ (it cannot cross $\mathbf{p}_1\mathbf{a}$ since $\check{\mathbf{x}}$ is visible, and it has to get very near to a point of the dotted circular arc $\mathbf{p}_1\check{\mathbf{a}}_\ell$ in Fig. 7 but not too close to $\mathbf{p}_1$ or $\check{\mathbf{a}}_\ell$). This contradicts Lemma 5.3(iii). $\square$

Corollary 5.5 *As in Lemma 5.4, let $\mathbf{a} \in \partial S_1$ and let $\check{\mathbf{a}}_\ell \in \partial T_2$ be the left nearest point of $\mathbf{a}$. Then for every neighborhood V of $\check{\mathbf{a}}_\ell$ there is a neighborhood U of $\mathbf{a}$ such that if $\check{C}$ denotes the portion of ∂T_2 lying in V and right of $\check{\mathbf{a}}_\ell$, and if we let $C = \mathrm{bisect}(\mathbf{p}_1, \check{C})$, then $\mathbf{a} \in C$ and the portion of ∂S_1 lying left of $\mathbf{a}$ and in U coincides with the portion of C lying left of $\mathbf{a}$ and in U.*

Proof. By Lemma 5.4 we know that the considered piece of ∂S_1 is contained in the bisector. The bisector bisect$(\mathbf{p}_1, \check{C})$ is the boundary of the convex set dom$(\mathbf{p}_1, \check{C})$ ([1], Lemma 3(iii)), and so it has to coincide with the considered piece of ∂S_1 (i.e., there cannot be any points of C in U and not in ∂S_1, since C is a convex curve). $\square$

Proof of Proposition 5.2. For contradiction let us assume $\mathbf{S} \neq \mathbf{T}$. We call $\mathbf{x}$ a *point of non-uniqueness* if $\mathbf{x} \in S_i \triangle T_i$, where $\triangle$ denotes symmetric difference.

Let

$$r_0 = \inf\{d(\mathbf{p}_i, \mathbf{x}) : \mathbf{x} \in S_i \triangle T_i, i = 1, 2, \ldots, n\}$$

be the infimum of distances of points of non-uniqueness to their respective sites. We have $r_0 > 0$, since each S_i and T_i contains a disk of nonzero radius around $\mathbf{p}_i$.

We note that there is no non-uniqueness at r_0 itself; that is, $S_i \cap B(\mathbf{p}_i, r_0) = T_i \cap B(\mathbf{p}_i, r_0)$ for all i, where $B(\mathbf{p}, r)$ denotes the disk of radius r centered at $\mathbf{p}$. This is because any closed convex set in $\mathbb{R}^d$ with nonempty interior equals the closure of its interior (and we apply this to $S_i \cap B(\mathbf{p}_i, r_0)$ and $T_i \cap B(\mathbf{p}_i, r_0)$).

Clearly, there is (at least one) index i that "causes" r_0; that is, with $r_0 = d(\mathbf{p}_i, S_i \triangle T_i)$. For notational convenience we assume that $i = 1$ is such. By a simple compactness argument, we can choose a sequence $(\mathbf{x}_j)_{j=1}^{\infty}$ of points in $S_1 \triangle T_1$ with $\lim_{j\to\infty} d(\mathbf{x}, \mathbf{p}_1) = r_0$ and such that the $\mathbf{x}_j$'s converges to a point $\mathbf{a}$. For convenience we assume that all the $\mathbf{x}_j$ lie left of $\mathbf{a}$ when viewed from $\mathbf{p}_1$.

By possibly exchanging the roles of $\mathbf{S}$ and $\mathbf{T}$, we may assume $\mathbf{x}_j \in T_1 \setminus S_1$for all j. Then $\mathbf{a}$ is a boundary point of S_1 since, on the one hand, it is in T_1 and T_1 coincides with S_1 up until radius r_0, and on the other hand, it is in the closure of the complement of S_1.

Let $\check{\mathbf{a}} \in \bigcup_{i=2}^{n} T_i$ be a nearest point of $\mathbf{a}$ and let $\check{\mathbf{a}} \in T_{i_0}$. By Lemma 5.3(i) we have $r_0 = d(\mathbf{p}_1, \mathbf{a}) = d(\mathbf{a}, \check{\mathbf{a}}) \geq d(\check{\mathbf{a}}, \mathbf{p}_{i_0})$. Thus $\check{\mathbf{a}} \in S_{i_0}$ as well, and it follows that $\mathbf{a}$ is a boundary point of T_1, too (since $T_1 \subseteq \mathrm{dom}(\mathbf{p}_1, \{\check{\mathbf{a}}\})$).

It follows that the set of nearest points of $\mathbf{a}$ in $\bigcup_{i=2}^{n} T_i$ coincides with the set of nearest points of $\mathbf{a}$ in $\bigcup_{i=2}^{n} S_i$. Let $\check{\mathbf{a}}_\ell$ be the left nearest point of $\mathbf{a}$. We fix notation so that $\check{\mathbf{a}}_\ell \in \partial T_2$ (then $\check{\mathbf{a}}_\ell \in \partial S_2$ as well).

By Corollary 5.4, a small portion of ∂S_1 left of $\mathbf{a}$ is uniquely determined by a small portion of ∂T_2 right of $\check{\mathbf{a}}_\ell$, and similarly for T_1 and S_2. Hence by the non-uniqueness assumption, ∂S_2 and ∂T_2 cannot coincide on any small neighborhood right of $\check{\mathbf{a}}_\ell$.

We will distinguish several cases. First, if $d(\check{\mathbf{a}}_\ell, \mathbf{p}_2) < d(\mathbf{p}_1, \mathbf{a}) = r_0$, then also a small neighborhood of $\check{\mathbf{a}}_\ell$ has distance to $\mathbf{p}_2$ smaller than r_0, and hence T_2 and S_2 coincide near $\check{\mathbf{a}}_\ell$, which is a contradiction. From now on we thus assume $d(\mathbf{p}_2, \check{\mathbf{a}}_\ell) = r_0 = d(\mathbf{a}, \check{\mathbf{a}}_\ell)$.

Next, we consider the angle $\alpha_\ell = \angle \check{\mathbf{a}}_\ell \mathbf{a} \mathbf{p}_1$; see Fig. 8. We claim that $\alpha_\ell \geq \pi$. Indeed, we have $S_1, T_1 \subseteq \mathrm{dom}(\mathbf{p}_1, \{\check{\mathbf{a}}_\ell\})$, and if $\alpha < \pi$, then this condition forces ∂S_1 and ∂T_1 in a small left neighborhood of $\mathbf{a}$ to be at distance smaller than r_0 to $\mathbf{p}_1$, which contradicts the assumed non-uniqueness.

We also need to consider the angle $\check{\alpha} = \angle \mathbf{p}_2 \check{\mathbf{a}}_\ell \mathbf{a}$; see Fig. 9. If $\check{\alpha} < \pi$, then by the same argument as above, small portions of ∂T_2 and ∂S_2 right of $\check{\mathbf{a}}_\ell$ coincide (since $T_2, S_2 \subseteq \mathrm{dom}(\mathbf{p}_2, \{\mathbf{a}\})$ etc.), which is a contradiction. Hence $\check{\alpha} \geq \pi$.

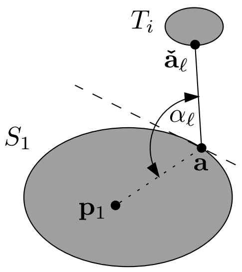

Figure 8: The case $\alpha_\ell < \pi$.

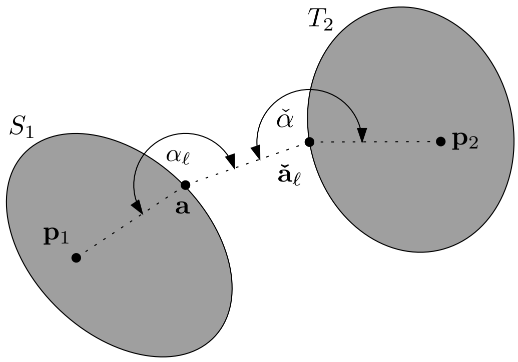

Figure 9: The angle $\breve{\alpha}$

We now deal with the case $\breve{\alpha} > \pi$. Here T_2 and S_2 are contained in the region Q that is the intersection of the halfplane $\mathrm{dom}(\mathbf{p}_2, \{\mathbf{a}\})$ with the halfplane h with boundary passing through $\breve{\mathbf{a}}_\ell$ and perpendicular to $\mathbf{a}\breve{\mathbf{a}}_\ell$; see Fig. 10. Clearly, for any point $\mathbf{x}$ in the dark gray wedge in Fig. 10, the nearest point in Q is $\breve{\mathbf{a}}_\ell$. Therefore, small portions of ∂S_1 and ∂T_1 left of $\mathbf{a}$ are contained in the bisector $\mathrm{bisect}(\mathbf{p}_1, \{\breve{\mathbf{a}}_\ell\})$ (a straight line), and in particular, they coincide—a contradiction finishing the case $\breve{\alpha} > \pi$.

Now we thus assume $\breve{\alpha} = \pi$. In order to proceed, we repeat for T_2, S_2 and $\breve{\mathbf{a}}_\ell$ some of the considerations made above for S_1, T_1, and $\mathbf{a}$, with left changed

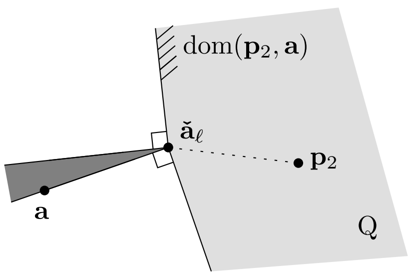

Figure 10: The region Q.

to right. First we can see that $\mathbf{a}$ is the right nearest point of $\breve{\mathbf{a}}_\ell$, for otherwise, small pieces of ∂T_2 and ∂S_2 right of $\breve{\mathbf{a}}_\ell$ would have distance at most r_0 to $\mathbf{p}_2$ and they would thus coincide there, a contradiction.

Second, we can get a contradiction as above if $\alpha_\ell > \pi$: For a small piece of ∂T_2 and ∂S_2 right of $\breve{\mathbf{a}}_\ell$, the nearest point is $\mathbf{a}$, hence these pieces would be the same straight segment.

Finally, we are left with the situation where $\alpha_\ell = \breve{\alpha} = \pi$ (in other words, $\mathbf{p}_1$, $\mathbf{a}$, $\breve{\mathbf{a}}_\ell$, and $\mathbf{p}_2$ are collinear), $\breve{\mathbf{a}}_\ell$ is the left nearest point of $\mathbf{a}$, and $\mathbf{a}$ is the right nearest point of $\breve{\mathbf{a}}_\ell$. Let Σ be a sufficiently narrow strip with one side given by the line $\mathbf{p}_1\mathbf{p}_2$ and the other side on the left of $\mathbf{a}$, let C_1 be the component of $\Sigma \cap \partial S_1$ adjacent to $\mathbf{a}$, and similarly, let C_2 be the component of $\Sigma \cap \partial T_2$ adjacent to $\breve{\mathbf{a}}_\ell$. Then C_1 and C_2 satisfy

$$C_1 = \operatorname{bisect}(\mathbf{p}_1, C_2) \cap \Sigma, \qquad C_2 = \operatorname{bisect}(\mathbf{p}_2, C_1) \cap \Sigma.$$

By the results of [1] (a small modification of Proposition 6, with the symmetric interval $(-a, a)$ replaced by $[0, a)$, and with almost no change in the proof), C_1 and C_2 are determined uniquely by these conditions; they are the "distance trisector curves" investigated in [1]. Since we have the same property for the appropriate pieces of ∂T_1 and ∂S_2, we again get a contradiction to the assumption that S_1 and T_1 should differ in every neighborhood of $\mathbf{a}$. Proposition 5.2 is proved. $\square$

6 Concluding Remarks

Non-local influence and sensitivity in zone diagrams. We sketch an interesting example. The left picture in Fig. 11 shows a zone diagram, a "dagger," with one isolated site $\mathbf{q}$ and three "flowers" marked by empty circles, where each flower has 6 sites arranged at the vertices of a tiny regular hexagon. As was observed in Section 2, if the flowers are very small, the region of the isolated site is close to the classical Voronoi region of $\mathbf{q}$. In the present case it is (almost) a skinny triangle. In the right picture, we have a small horizontal dagger on the top. Then there are two tiny flowers and an isolated site on the right, and these flowers plus the tip of the small horizontal dagger induce a region of the isolated site, which is also (almost) a skinny triangle. This makes a larger vertical dagger on the right. This can be iterated in a spiral-like fashion with any number of progressively larger daggers. (Of course, a formal proof that the regions truly look as claimed would be longer.)

This daggers example witnesses two things. First, the location of the tip of the last (largest) dagger depends on the location of *all* of the flowers and previous isolated sites. Therefore, zone diagrams posses no locality in a sense similar to classical Voronoi diagrams. Second, changing the location of one of the flowers in the first (smallest) dagger has a large influence on the position of the tip of that dagger, which in turn exerts an even much larger influence on the tip of the second dagger, and so on. We have a complicated "leverage effect," again quite unlike in classical Voronoi diagrams. We can also see that

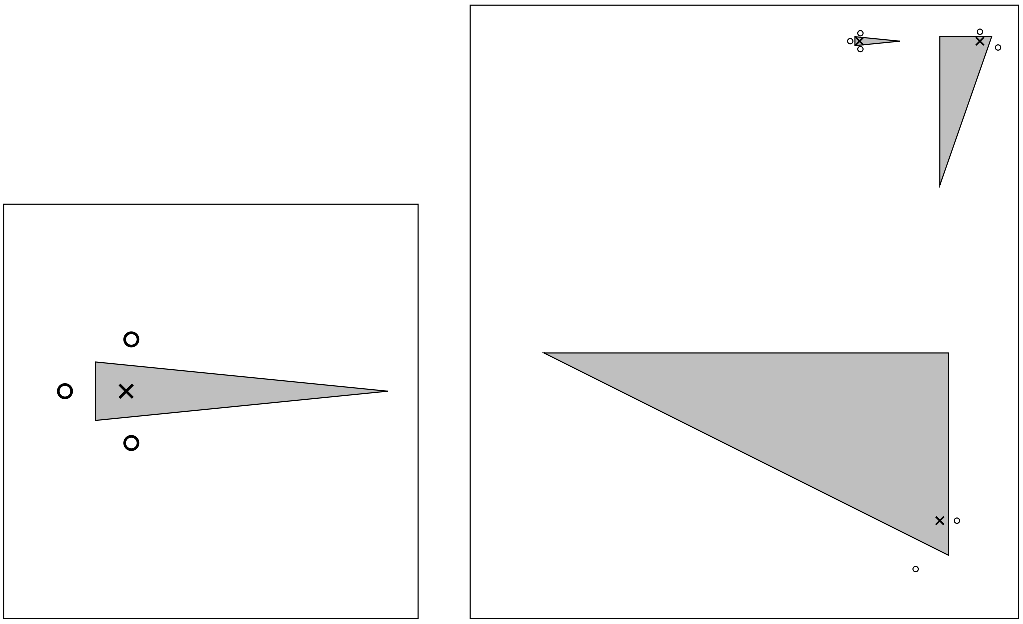

Figure 11: The "daggers" example, documenting the long-range influence in zone diagrams and the great sensitivity to small changes in site positions. Crosses represent sites and small circles represent tiny flowers.

the convergence of the iterative algorithm from Section 1 is likely to be relatively slow on this example: The tip of the first dagger has to stabilize very precisely before the second dagger has a chance to approach its final state, etc. Therefore, a bound on the convergence rate must take the number of sites into account, as well as some parameter like the ratio of the maximum and minimum site distances.

Combinatorial complexity of zone diagrams. For a zone diagram $\mathbf{R}$, *singular points* on the boundary of a region R_i are those with at least two distinct nearest points in $\bigcup_{j\neq i} R_j$ (this notion has been considered in Section 5). We could regard the singular points as an analog of Voronoi vertices in a classical Voronoi diagram, and the segments of ∂R_i between consecutive singular points as an analog of Voronoi edges.

It is not hard to show that each of these Voronoi edges e is contained in the bisector of $\mathbf{p}_i$ and some R_j, $j \neq i$, which easily implies that e is of class C^1 (with continuous first derivative). We can prove that the number of singular points, as well as the number of "Voronoi edges," is $O(n)$, where n is the number of sites (we intend to publish a proof elsewhere). However, it should be noted that a single "Voronoi edge" $e \subseteq \partial R_i$ can still be complicated: We have $e \subseteq \mathrm{bisect}(\mathbf{p}_i, e')$, where e' is a piece of ∂R_j (for some j) which may contain many singular points.

Crystal growth: Intuition and alternative algorithm. Here is our original intuition for the uniqueness proof. We imagine that a crystal starts growing from each site $\mathbf{p}_i$ at time $t = 0$, and we let $R_i(t)$ denote its shape at time $t \geq 0$. Initially each of the crystals grows everywhere along its boundary at unit speed, but as soon as the distance of a boundary point $\mathbf{x} \in \partial R_i(t)$ to some $R_j(t)$ becomes $d(\mathbf{x}, \mathbf{p}_i)$, the growth at $\mathbf{x}$ stops. It seems intuitively clear that the result of this growth process should be a zone diagram, and actually the only possible zone diagram. But proving it seems to require some kind of "induction on the radius," and here the usual troubles with the continuous nature of the reals start (resembling the troubles with the intuitively obvious arguments of the old masters of calculus, arguments which were later replaced by the much more complicated-looking proofs in contemporary textbooks of analysis).

Our uniqueness proof shows that given $\mathbf{R}(t) = (R_1(t), R_2(t), \ldots, R_n(t))$ at some time t, we can uniquely extend it to $\mathbf{R}(t + \varepsilon)$ for some $\varepsilon = \varepsilon(t) > 0$, and this is even "efficient" in the sense that the new pieces of the boundary are given as bisectors of sites and old pieces, or as pieces of the distance trisector curve. We could thus start with $\mathbf{R}(t_0)$ for a suitable t_0 where all $R_i(t)$ are disks of radius t_0, extend to t_1, then to t_2, etc., but if we take the proof as is, the steps $t_{k+1} - t_k$ might possibly get smaller and smaller and we might get an infinite but *bounded* sequence $t_0 < t_1 < t_2 < \cdots$.

It turns out that we can get away with finitely (and even polynomially) many time steps if we are willing to make the computation of a bisector of an already computed curve and a site in a single step (as well as the computation of the distance trisector curve). But such operations may be too complex to be considered as reasonable computational primitives, and further work is still needed.

Acknowledgments

We would like to thank Tomáš Kaiser for useful suggestions and discussions concerning fixed-point theorems and Eva Kopecká for suggesting an alternative (and simpler) approach to the existence proof.

References

[1] T. Asano, J. Matoušek, T. Tokuyama. The distance trisector curve, *Advances in Mathematics*, in press. Extended abstract in *Proc. 38th ACM Symposium on Theory of Computing*, 2006.

[2] T. Asano and T. Tokuyama. Drawing Equally-Spaced Curves between Two Points, *Proc. Fall Conference on Computational Geometry,* Boston, Massachusetts, November 2004, pages 24–25.

[3] F. Aurenhammer. Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure. *ACM Computing Surveys* 23,3(1991) 345–405.

[4] D. W. Curtis and R. M. Schori. Hyperspaces of Peano continua are Hilbert cubes. *Fundam. Math.* 101(1978), 19–38.

[5] A. Okabe, B. Boots, K. Sugihara. *Spatial Tessellations, Concepts and Applications of Voronoi Diagrams*, John Wiley & Sons, New York, NY 1992.

[6] E. Zeidler, *Nonlinear functional analysis and its applications. Volume I: Fixed-point theorems*, 2nd corr. printing, Springer, New York, 1993.

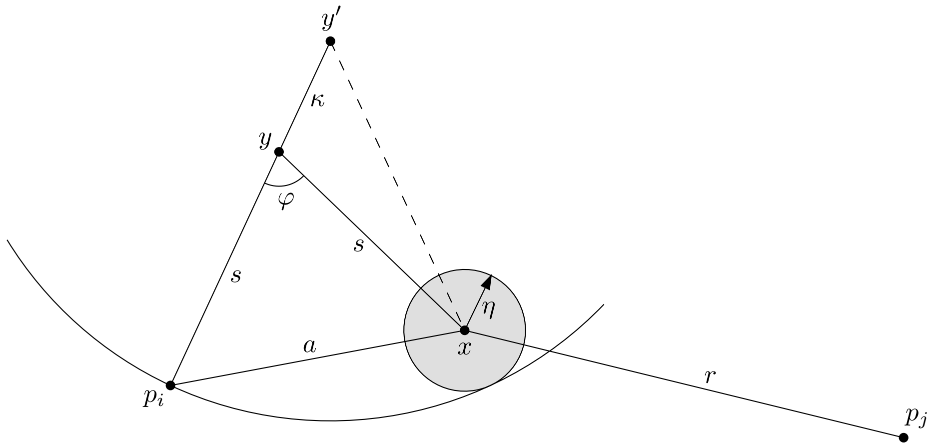

Figure 12: Illustration to the proof of continuity of F.

7 Appendix

7.1 Proof of Lemma 4.4

We will actually prove that F is C-Lipschitz for a (large) constant C depending on the point set P. Let $\boldsymbol{\rho}, \boldsymbol{\rho}' \in K$. Let us put $\boldsymbol{\sigma} = F(\boldsymbol{\rho})$, $\boldsymbol{\sigma}' = F(\boldsymbol{\rho}')$, $\delta = \|\boldsymbol{\rho} - \boldsymbol{\rho}'\|_\infty$, $\varepsilon = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}'\|_\infty$. To prove continuity, we want to show a strictly positive lower bound on δ for every $\varepsilon > 0$. (It seems that apriori we cannot assume ε small.)

Let ε be attained for i and β; that is, $\sigma_i'(\beta) - \sigma_i(\beta) = \varepsilon$ (should the sign be opposite, we interchange $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$). Let $\mathbf{y}$ be the boundary point of $S_i = \mathrm{reg}_{\mathbf{p}_i}(\sigma_i)$ in direction β, and let $s = |\mathbf{p}_i\mathbf{y}| = \tan(\sigma_i(\beta))$. See Fig. 12.

Let us set $s' = \tan(\sigma_i(\beta) + \varepsilon/4)$, and let $\mathbf{y}'$ be the point of $S_i' = \mathrm{reg}_{\mathbf{p}_i}(\sigma_i')$ at distance s' from $\mathbf{p}_i$ in direction β. This choice, instead of $\mathbf{y}'$ lying on the boundary of S_i' (which looks more natural), guarantees two things: First, $\mathbf{y}'$ is at finite distance from $\mathbf{p}_i$, and second, we have $\kappa = s' - s \le s$. To verify the latter claim, we note that we may assume $\varepsilon = \frac{\pi}{2} - \arctan s$, and we use the Mean Value Theorem to bound $s' - s = \tan(\arctan s + \varepsilon/4) - \tan(\arctan s) = \frac{\varepsilon}{4} \cdot \frac{1}{\cos^2(\arctan s + \varepsilon/4)} = \frac{\varepsilon}{4} \cdot \frac{1}{\sin^2(3\varepsilon/4)}$. Now $\varepsilon \le \pi/2 - \arctan 1 = \pi/4$, and since for $0 \le x \le \frac{3}{16}\pi$ we have $\sin x \ge 0.9x$, we obtain $s' - s \le \frac{\varepsilon}{4} \cdot \frac{1}{(0.9 \cdot 0.75\varepsilon)^2} \le \frac{0.6}{\varepsilon}$. On the other hand, $s = \tan(\frac{\pi}{2} - \varepsilon) = \frac{1}{\tan \varepsilon} \ge \frac{\pi}{4\varepsilon} > \frac{0.6}{\varepsilon}$, and $s' - s \le s$ is proved.

Since $\mathbf{y}$ is a boundary point of S_i, it is easy to see that there has to be a boundary point $\mathbf{x}$ of some R_j, $j \ne i$ (where $(R_1, \ldots, R_n) = \mathbf{reg}(\boldsymbol{\rho})$), such that $|\mathbf{xy}| = |\mathbf{yp}_i| = s$ (briefly, points arbitrarily close to $\mathbf{y}$ but outside S have points of some R_j at distance arbitrarily close to s, and a limit argument using compactness provides the desired $\mathbf{y}$). On the other hand, the open disk of radius $s' = s + \kappa$ centered at $\mathbf{y}'$ is disjoint from all R_k', $k \ne i$, for otherwise, $\mathbf{y}'$ wouldn't lie in S_i'.

If we prove some lower bound η on the difference $s' - |\mathbf{y}'\mathbf{x}|$, then the open disk of radius η centered at $\mathbf{x}$ is disjoint from all R_j', and in particular, the point $\mathbf{x}'$ on the segment $\mathbf{p}_j\mathbf{x}$ lying at distance η from $\mathbf{x}$ cannot be inside R'_j. It follows that

$$\delta = \|\boldsymbol{\rho} - \boldsymbol{\rho}'\|_\infty \geq \arctan(r) - \arctan(r-\eta) \geq \frac{\eta}{1+r^2} \geq \frac{\eta}{2r^2}$$

by the Mean Value Theorem (we have $r - \eta \geq 1$ since R'_j contains the unit disk centered at $\mathbf{p}_j$).

To estimate η, we consider the triangle $\triangle\mathbf{y}'\mathbf{y}\mathbf{x}$, and by the Cosine Theorem we obtain

$$\begin{aligned}
|y'x| &= \sqrt{s^2+\kappa^2-2s\kappa\cos(\pi-\varphi)} \\
&= \sqrt{s^2+\kappa^2+2s\kappa\cos\varphi} \\
&= s'\sqrt{1-\frac{2s\kappa}{s'^2}(1-\cos\varphi)} \\
&\leq s'\left(1-\frac{s\kappa}{s'^2}(1-\cos\varphi)\right) \\
&= s'-\frac{s\kappa}{s'}(1-\cos\varphi),
\end{aligned}$$

where we have used $\sqrt{1-z} \leq 1 - z/2$ for $0 < z < 1$. The Cosine Theorem for the triangle $\triangle\mathbf{p}_i\mathbf{y}\mathbf{x}$ then yields

$$a^2 = |\mathbf{p}_i\mathbf{x}|^2 = 2s^2(1-\cos\varphi),$$

and altogether we have

$$\eta = s' - |\mathbf{y}'\mathbf{x}| \geq \frac{a^2\kappa}{2ss'} \geq \frac{a^2}{4s^2}\kappa$$

(using the inequality $s' \leq 2s$ mentioned above).

For a we use the obvious estimate $a \geq 1$ (from the fact that S_i contains the unit disk centered at $\mathbf{p}_i$), as well as $a = |\mathbf{p}_i\mathbf{y}| \geq |\mathbf{y}\mathbf{p}_j| - |\mathbf{p}_i\mathbf{p}_j| \geq r - \Delta$, where Δ denotes the diameter of P. Together we have $a \geq \max(1, r-\Delta) \geq r/2\Delta$ (distinguishing the cases $r \leq 2\Delta$ and $r \geq 2\Delta$). Finally, for $\kappa = s' - s$ we have $\arctan(s') - \arctan(s) = \varepsilon/4$, and the Mean Value Theorem (as usual) gives $\kappa = s' - s \geq \frac{1}{4}\varepsilon \cdot (1+s^2) \geq \frac{1}{4}\varepsilon s^2$.

Putting the chain of inequalities together, we have

$$\|\boldsymbol{\rho} - \boldsymbol{\rho}'\|_\infty \;=\; \delta \geq \frac{\eta}{2r^2} \geq \frac{a^2}{4s^2 \cdot 2r^2} \cdot \frac{\varepsilon}{4}s^2 \geq \frac{r^2}{4\Delta^2 \cdot 8r^2} \cdot \frac{\varepsilon}{4} = \frac{\varepsilon}{128\Delta^2}.$$

This shows that **Dom** is continuous, and even C-Lipschitz for a suitable constant C. $\square$