Conditional Facility Location Problems with Continuous Demand and a Polygonal Barrier

Conditional Facility Location Problems with Continuous Demand and a Polygonal Barrier

Thomas Byrne Jörg Kalcsics
School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK

Abstract

We consider facility location problems where n facilities are present in a convex polygon in the rectilinear plane, over which continuous and uniform demand is distributed and within which a convex polygonal barrier is located (removing all demand and preventing all travel within the barrier), and the optimal location for an additional facility is sought. We start with an in-depth analysis of the representation of the bisectors of two facilities affected by the barrier and how it is affected by the position of the additional facility. Following this, a detailed investigation into the changes in the structure of the Voronoi diagram caused by the movement of this additional facility, which governs the form of the objective function for numerous facility location problems, yields a set of linear constraints for a general convex barrier that partitions the market space into a finite number of regions within which the exact solution can be found in polynomial time. This allows us to formulate a polynomial exact algorithm that makes use of a triangular decomposition of the incremental Voronoi diagram and the first order optimality conditions.

Keywords – location, planar facility location, spatial demand, Voronoi diagrams, barriers

1 Introduction

The task of determining optimal locations for a set of facilities is of strategic importance and so, naturally, a vast number of models and algorithms have appeared in the literature ever since. It all started with the seminal works of Launhardt (1900) and Weber (1909), who were the first to study the problem in an industrial context (finding an optimal location for a plant on the plane in the presence of a single market and two suppliers); see Wesolowsky (1993) for a good discussion. In the 1960s, location problems in other domains evolved, namely on networks (Hakimi, 1964) and discrete location problems (Manne, 1964). But ever since the works of Launhardt and Weber, planar location problems have received considerable and ongoing attention in the literature: see the books of Drezner (1995a); Drezner and Hamacher (2002); Laporte, Nickel, and Saldanha da Gama (2019).

In this paper we will consider five well-known location problems (though the methodology is applicable to a range of continuous demand location problems): the median and antimedian problems, the centre problem, the market share problem, and the maximal covering problem. The objectives are: in the median (antimedian) problem, to minimise (maximise) the combined distance that all customers travel; in the centre problem, to minimise the maximum distance that any customer travels; in the market share problem, to maximise the total demand captured by a selected company's facilities (supposing that the selected company owns the new facility and the others belong to the company or its competitors); and in the maximal covering problem, to maximise the demand of all customers captured within a particular coverage radius from a facility.

The vast majority of planar location problems rely on the assumptions that the demand of customers is represented by a finite set of discrete points and that the placement of facilities on the plane as well as travel across the plane is not restricted, as is evidenced by the three abovementioned books.

In an industrial context, as laid out by Launhardt and Weber, a discrete set of demand points often adequately reflects the problem. For commercial and public service facilities in an urban environment,

however, there can be millions of potential customers, and it is impracticable to represent every customer site as a separate demand point. One option to deal with such a situation is to aggregate customers into a smaller number of meta-customers, e.g. by postal codes, census tracts, or wards. This, however, introduces various kinds of aggregation errors to the problem, which can be quite pronounced (Drezner, 1995b; Francis & Lowe, 2019). A second, more accurate approach is to model customer demand as continuously distributed over the specified area. Moreover, the uncertain and sporadic nature of some demands lends itself much better to a continuous distribution over the region under consideration than to a discrete one. The first treatment of a location problem with continuous demand (and Euclidean distances) is due to Lösch (1954). Since then, a significant amount of research has been dedicated to planar location problems with continuous demand distributions: see Newell (1973); Erlenkotter (1989); Fekete, Mitchell, and Beurer (2005); Averbakh, Berman, Kalcsics, and Krass (2015) and references therein.

Concerning the second assumption, the region of demand and the region over which a facility can be feasibly located are often assumed to be the whole plane or, at least, convex polygons. However, this supposition is not a realistic one for many real world applications, as there are often obstacles in cities which restrict the placement of facilities as well as travel, for example, rivers, lakes, parks, hills, or highways and rail tracks. Location problems containing forbidden areas that can be traversed, but where placement is not allowed, have been studied, e.g. in Aneja and Parlar (1994); Hamacher and Nickel (1995). If the traversal of these areas is also not allowed, then we can no longer use 'straight-line' distances and must instead revert to geodesic distances; these areas are usually called barriers. One of the first works on barrier location problems is Katz and Cooper (1981), who consider a circular barrier. Since then, barrier location problems have received ongoing attention in the literature: see Larson and Sadiq (1983); Aneja and Parlar (1994); Klamroth (2001); Bischoff and Klamroth (2007); Canbolat and Wesolowsky (2012); Oğuz, Bektaş, and Bennell (2018) and references therein.

However, to the best of our knowledge, so far continuous demand and barrier problems have only been studied independently of one another, and this is the first work that combines both of them into a single model. The added value of combining both is immediately apparent since the setting in which spatially distributed demand is most appropriate, namely in an urban environment, is also the one where barriers have the biggest impact on facility placement and, especially, on travelling. A current example, owing to the COVID-19 pandemic, is optimally locating roadside testing facilities to reduce the spread of infection and save lives: such facilities seek to serve a large population (requiring continuous demand to model) and customers (patients) travel (approximately according to rectilinear distance) to the facilities within an urban area full of barriers. Given its urban environment, this is a scenario which critically requires an algorithm dealing with continuous demand, rectilinear distances, and barriers.

Solving either of the two extensions independently is challenging, but both combined is even more so. Moving from discrete demand points to a continuous demand distribution significantly changes the structure of the problem. Planar location problems with discrete demand and Euclidean distances can be solved analytically, without having to investigate any geometric properties of the problem (Weiszfeld, 1937; Xue, Rosen, & Pardalos, 1996). For the Manhattan norm, one can even discretise the solution space and pose the problem as an equivalent discrete location problem (Francis & White, 1974). Not so for continuous demand distributions with multiple facilities. Following stipulation of the locations of all p facilities, the demand space has to be subdivided into areas known as 'Voronoi cells'; the resulting partition is called the 'Voronoi diagram'. Moreover, in order to solve the problem, merely calculating the Voronoi diagram based on the facility locations is insufficient. Instead, it is essential to study the *structure* of the Voronoi diagram, i.e. the position and the geometry of the cell boundaries, and how it changes dynamically when one or more of the facilities 'move around' on the plane. Unfortunately, this structure can alter dramatically, which makes it generally nigh impossible to represent the objective function of the problem in closed form, and renders the formulation of the underlying optimisation problems very challenging (Averbakh et al., 2015).

The following example gives an idea of these difficulties of representation (even before the addition of a barrier). Figure 1 shows the current facilities $A_1, \ldots, A_4$ and their Voronoi cells representing the trading regions of these facilities within the market space $\mathcal{P}$ along with a barrier $\mathcal{B}$. Regarding

the two potential locations Z and Z', we notice that their Voronoi cells have significantly different shapes despite the proximity of the trial facilities. Hence, computing the demand for the market share problem (or any other objective function) over these Voronoi cells would result in functional expressions that differ considerably, making it very challenging to optimise the location of Z over the market region.

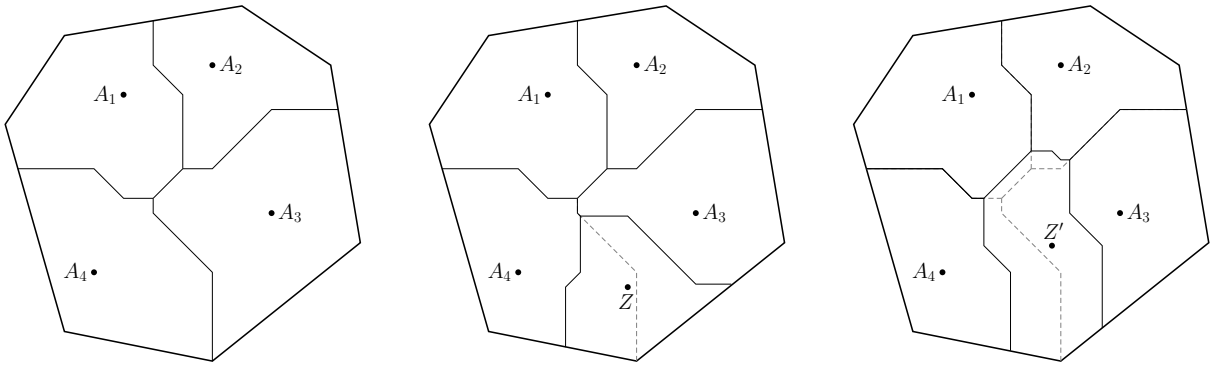
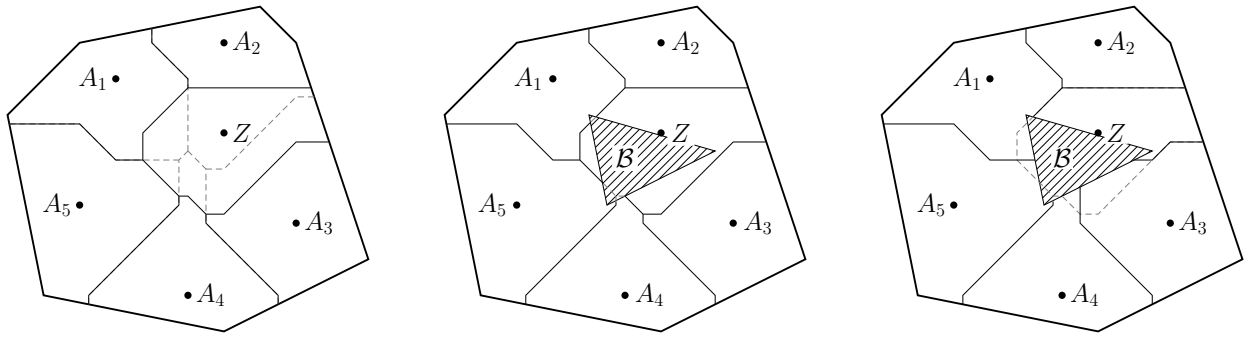


Figure 1: The market areas of four facilities within a polygonal market region and changes when a new facility Z or Z' is added.

These difficulties are exacerbated when we introduce barriers. In this case, we have to replace the regular 'straight-line' distances with their geodesic counterparts. While some solution methods for barrier location problems with discrete demand require the calculation of the barrier-restricted Voronoi diagram, this is always the Voronoi diagram of the set of fixed demand points, and thus is static and not affected by the locations of the facilities (Dearing, Hamacher, & Klamroth, 2002; Dearing, Klamroth, & Segars, 2005). Moreover, since we are working with continuous demand we cannot even derive such Voronoi diagrams based on discrete customer locations in the first place. This is a stark contrast to models for continuous customer demand and invalidates any exact solution method for discrete demand barrier problems.

To empirically demonstrate the additional difficulties we are facing due to the barrier, we present an example where we are given five existing locations for facilities and we want to find an optimal location for a sixth facility, with the goal of maximising the area of its Voronoi cell, i.e. maximising its market share for a uniform demand distribution. Figure 2a shows the five existing facilities A_1, ..., A_5 within a convex polygonal area and what could well be the optimal location, Z. We show the parts of the Voronoi diagram of the five existing facilities that disappeared after the addition of Z in dashed grey.



(a) A solution to a barrier-free facility location problem.

(b) The barrier-free solution in the presence of a barrier.

(c) The equivalent barrier-restricted solution.

Figure 2: Demonstration of some drawbacks of sharing solutions between barrier-free and barrier-restricted location problems.

Moreover, if the real-life situation happened to include barriers which were not included in the modelling, then a multitude of shortcomings could befall the valuation of this location Z. A simple one is that Z may not be feasible in the barrier-restricted problem, because we cannot locate in the barrier and Z may lie within the barrier. Even if Z is still feasible, the barrier can severely affect the originally calculated captured demand, because the area captured may in reality be much less than promised in the barrier-free case if the barrier swallows up much of the Voronoi cell of Z. We can see in Figure 2b that a large proportion of what we were previously awarding to Z is swallowed by $\mathcal{B}$, so the captured demand in the original Voronoi cell of Z is much less in the barrier-restricted problem. Furthermore, we can see that the barrier has cut off the leftmost area of the original Voronoi cell of Z. It no longer makes any sense to assign this isolated area to Z. Additionally, the barrier now enforces that, in order to access the bottommost area of the original Voronoi cell of Z, one must travel around the rightmost vertex of $\mathcal{B}$. Because the barrier warps distance, it is nonsensical to still assign the area within the barrier-free Voronoi cell of Z to Z in the barrier-restricted problem. The barrier can greatly influence the Voronoi diagram and Figure 2c shows that the barrier-restricted cell of the barrier-free optimum can easily be drastically different, and lead to a particularly poor solution. These shortcomings befall not only the market share problem, but any such facility location problem.

The ideas here become increasingly prevalent when one considers the fact that a barrier-free modelling of a barrier-restricted problem will contain a large area wherein facilities are not located (because of the barrier) and so optimal solutions of the barrier-free problem may be even more likely to be placed in or near the barrier and so be very significantly affected by the barrier's reintroduction.

In this paper, we will explore the five previously mentioned location problems over a convex polygonal market region in the rectilinear plane containing one known convex polygonal barrier outside of which some facilities are already fixed within the space. Moreover, we assume that the demand is continuous and uniform, with customers being served by the nearest facility only. We concentrate on conditional location problems where we wish to find the optimal location for an additional facility, the location of which has no restriction within the traversable polygon.

Our goal is to develop an algorithm to work with this continuous demand and an arbitrary convex polygonal barrier. To that end, in this paper we derive the structural properties of geodesic Voronoi diagrams with rectilinear distances. This enables us to solve the overlying optimisation problem itself by restricting the location of the new facility to a sub-region where the resulting geodesic Voronoi diagram is 'structurally identical' for every point in the region. Given such regions, we derive a parametric representation of the objective function which is valid for any location in the region. By this means we optimise the location of the new facility over this region using classical non-linear programming techniques, and the best optimal location of each sub-region is the optimal solution to the problem.

A summary of relevant literature is provided in Section 2. Basic definitions and notations are then introduced in Section 3, followed by an analysis of the changes to the objective function's representation resulting from the movement of the additional facility and the calculation of partitioning lines to preserve the representation, allowing us to derive a partition as described above. In Section 5 we compute the objective functions for the five conditional location problems previously defined. Conclusions are in the last section, along with a description of current work and suggestions of applications relating to further research.

2 Literature Review

As alluded to in the introduction, there is some literature on facility location problems with continuous demand, but it is minuscule in comparison to that for discrete demand problems. The first discussion of location problems with continuous demand and rectilinear distances appears in Marucheck and Aly (1981). They proposed a branch-and-bound algorithm for a rectangular market region model over which demand is distributed uniformly. An unbounded market region with uniform demand is studied in the general optimal market area model in Erlenkotter (1989) with l_1, l_2, and block norms, which also includes facility costs. Different supply area shapes of each facility are trialled (circle, square,

diamond, and hexagon) and closed-form expressions for the optimal size of the supply area are derived, offering the opportunity to assess the sensitivity of non-optimally sized, or shaped, supply areas. All five problems we will explore are addressed in Averbakh et al. (2015) with uniform demand over a convex polygon with rectilinear distances. After identifying a partition of the market area which preserves the structural identity of the Voronoi diagram, they devise an exact polynomial algorithm for all five problems.

Exact polynomial algorithms solving the median problem for a single facility are proposed in Fekete et al. (2005) for straight-line and geodesic rectilinear distances and continuous demand. This is performed for polygonal market areas with and without holes; with holes it is shown to be NP-hard for multiple facilities whose number is part of the input. They assert that their approach is applicable to multifacility problems if the boundaries of the Voronoi regions can be suitably represented, however without giving any details. The two-facility problem is considered with l_1 and l_2 distances in Murat, Verter, and Laporte (2011). The problem is displayed as a two-dimensional boundary value problem with optimality conditions found, and is solved by means of a two-dimensional shooting algorithm.

Addressing the centre problem, Suzuki and Drezner (1996) derive upper and lower bounds for the p-centre problem for a convex polygonal area with uniform demand in Euclidean space by proposing a random-start heuristic based on Cooper's location-allocation algorithm. This is extended in Wei, Murray, and Xiao (2006) to cover non-convex demand regions with holes, wherein they apply their heuristic in order to find a preferable location for warning sirens.

With regard to market share problems, Okabe and Suzuki (1987) and Okabe and Aoyagi (1993) explore the equilibrium arrangements of competitive firms under a variety of premises (e.g. about the number of facilities or the market area's shape). To that end they run simulations involving techniques such as the gradient descent method where each facility is permitted to relocate once in each time interval.

Regarding maximal covering location problems, Matisziw and Murray (2009a) describe the extension of their earlier study on discs to arbitrarily-shaped demand spaces. Murray and Tong (2007) solve the problem with a demand defined by points, line segments, or polygons with Euclidean distance. Murray, O'Kelly, and Church (2008) suggest partitioning the demand space into regular polygons for location set covering problems, using the Euclidean metric in their warning siren study.

There have been few forays into facility location problems with barriers, and we are not aware that any have broached the continuous demand problem. Katz and Cooper (1981) study the Weber problem for a given a discrete set of demand points, a circular barrier, and Euclidean distances. They first show how to calculate the geodesic Euclidean distance between two points on the plane. Afterwards, to solve the problem, they propose to convert the non-linear constraint minimisation problem into a sequence of non-linear unconstrained minimisation problems. Aneja and Parlar (1994) consider the same setting, but with polygonal forbidden regions as well as with polygonal barriers. For the former, they show that the optimal solution either coincides with the optimal solution of the unconstrained problem, or it lies on the boundary of a forbidden region. For the latter, they propose a simulated annealing heuristic. Larson and Sadiq (1983) study the p-median problem with Manhattan distances and arbitrary barriers. Analysing the geometry of shortest geodesic paths, they identify a finite set of points on the plane that contains an optimal solution.

Given a discrete set of existing customers and convex polyhedral barriers, Klamroth (2001) finds the location of a facility to optimise any convex objective function in the distances between the facility and customers by partitioning the space into smaller subproblems. Dearing et al. (2002) discuss the 1-centre problem in l_1 for discrete demand and a set of convex polyhedral barriers. They show that the optimal location is found in the set of intersection points between barrier-restricted bisectors. This is extended in Dearing et al. (2005) to the case of general block distances. In a more recent paper, Oğuz et al. (2018) discretise not only the location space but also the shortest-path space. For the latter, they discretise the calculation of the distances between the customer demand points and the potential new facility, thereby reducing the traversable region to a network as opposed to continuous space. These discretisations transform the continuous multi-facility Weber problem into a discrete problem, which can then be formulated as a mixed-integer linear programming problem. Furthermore, both deterministic and probabilistic barriers are considered in this formulation.

Regarding heuristics, Bischoff and Klamroth (2007) design a genetic algorithm for the 1-Weber problem with discrete demand and convex polyhedral barriers in a general metric space, which determines iteratively the subproblems (as noted above) to solve.

Finally we refer the reader to Canbolat and Wesolowsky (2012) for a novel and hands-on approach to solving such discrete problems. In their paper the Varignon frame method is adapted to the barrier-restricted 1-Weber problem in l_2 with discrete demand and one convex barrier, using physical weights and strings to find the optimal solution mechanically. We encourage any keen reader with time to spare to enjoy recreating some of these spidery contraptions.

3 Definitions and Notations

Between any two points $P = (p_x, p_y)$ and $Q = (q_x, q_y)$ in $\mathbb{R}^2$, the rectilinear distance is $l_1(P,Q) = |p_x - q_x| + |p_y - q_y|$ and the bisector (the set of points that are equidistant from P and Q) is $B(P,Q) = \{X \in \mathbb{R}^2 \,|\, l_1(P,X) = l_1(Q,X)\}$. Let $B^{\leq}(P,Q) = \{X \in \mathbb{R}^2 \,|\, l_1(P,X) \leq l_1(Q,X)\}$. The relative positions of P and Q decide the bisector's shape. These positions are defined by the following three expressions: $|p_x - q_x| \diamond |p_y - q_y|$, $p_x \diamond q_x$, and $p_y \diamond q_y$ where $\diamond \in \{\leq, \geq\}$. For fixed P, fixing an inequality for each of the three relations gives a set of points Q called a configuration cone. The first one is defined as $\mathcal{CC}^1(P) := \{Q \in \mathbb{R}^2 \,|\, p_x \leq q_x, p_y \leq q_y, p_x - q_x \leq p_y - q_y\}$ with the other cones $\mathcal{CC}^2(P), \ldots, \mathcal{CC}^8(P)$ created analogously and labelled anticlockwise as displayed in Figure 3. The configuration lines $\mathcal{CL}^k(P) := \mathcal{CC}^k(P) \cap \mathcal{CC}^{k-1}(P), k = 1, \ldots, 8$ bound the configuration cones (where $\mathcal{CC}^0(P) := \mathcal{CC}^8(P)$).

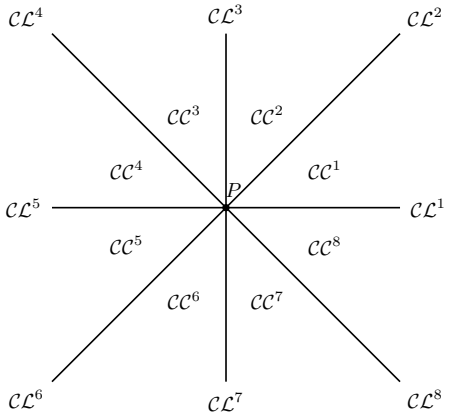


Figure 3: The eight configuration cones of P.

The bisector $B(P,Q)$ is piecewise linear with exactly three pieces if $Q \in \text{int}(\mathcal{CC}^k(P)))$ (where $\text{int}(\star)$ represents the interior of $\star$). If $Q \in \mathcal{CL}^k(P), k = 1, 3, 5, 7$, then the bisector is a straight line in which case we call it degenerate. If $Q \in \mathcal{CL}^k(P), k = 2, 4, 6, 8$, then the bisector consists of two quarter planes connected by a diagonal (note that a diagonal line is considered to be one with a gradient of ± 1). There exist four distinct shapes of non-degenerate bisectors for $p_y \leq q_y$, for Q in either $\mathcal{CC}^1(P)$, $\mathcal{CC}^2(P)$, $\mathcal{CC}^3(P)$, or $\mathcal{CC}^4(P)$, as displayed left to right in Figure 4. Identical plots are obtained for the other configuration cones with $p_y \geq q_y$, only with P and Q exchanged.

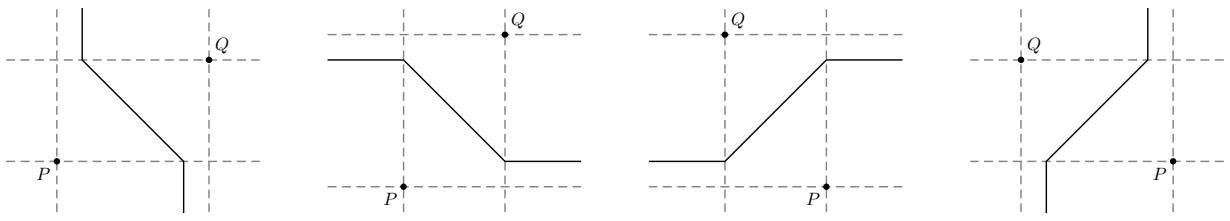


Figure 4: The four different shapes for non-degenerate bisectors for $p_y \leq q_y$.

Now we shall briefly introduce the theory of Voronoi diagrams; see Okabe, Boots, Sugihara, and Chiu (2000) for a more in-depth review. For a set $\{A_1, \ldots, A_n\} \in \mathbb{R}^2$ of $n \geq 2$ distinct points, define the Voronoi cell of A_i to be $V(A_i) = \bigcap_{j=1}^{n} B^{\leq}(A_i, A_j) = \{Q \in \mathbb{R}^2 \,|\, l_1(Q, A_i) \leq l_1(Q, A_j) \forall j \in \{1, \ldots, n\}\}$ and we say A_i is the generator of $V(A_i)$. The Voronoi diagram for $\{A_1, \ldots, A_n\}$ is denoted by $\mathcal{VD}(A_1, \ldots, A_n) = \{V(A_1), \ldots, V(A_n)\}$. The line between A_i and each vertex of the Voronoi cell is wholly contained inside the polygon, a critical property called star-shaped. We assume that no two generators $\{A_1, \ldots, A_n\}$ lie on the same diagonal line so as to circumvent the tribulation of their bisectors (a diagonal connecting two quarter planes). This assumption is not prohibitive in practice since, at least in this review, $\{A_1, \ldots, A_n\}$ represent the current facility locations, the coordinates of which can be slightly perturbed (due to a measurement error, say).

Moreover, we only consider the Voronoi diagram within the market region, $\mathcal{P}$. Let $\mathcal{P} \subset \mathbb{R}^2$ be a compact convex m-sided polygon and suppose $A_1, \ldots, A_n \in \mathcal{P}$. We redefine the Voronoi cells $V(A_i) = V(A_i) \cap \mathcal{P}$ which, since $\mathcal{P}$ is convex, are still star-shaped.

We will consider uniform demand over $\mathcal{P}$ and assume that each demand point is served by the closest facility. We will now introduce a convex p-sided polygonal barrier $\mathcal{B}$ with vertices $B_i = (b_{i_x}, b_{i_y})$ within which there is zero demand, and across which travel is not permitted. Now that a barrier exists, the distance between P and Q is no longer the rectilinear distance l_1, but the distance of the rectilinear path from P to Q around the barrier, say $l_1^{\mathcal{B}}(P, Q)$. Since the barrier alters the measuring of distances between points, it can affect the bisectors and therefore the Voronoi diagram. The new Voronoi cell of A_i, $V(A_i) = \{Q \in \mathcal{P} \setminus \text{int}(\mathcal{B}) \,|\, l_1^{\mathcal{B}}(Q, A_i) \leq l_1^{\mathcal{B}}(Q, A_j) \forall j \in \{1, \ldots, n\}\}$ can have a very different shape to the original Voronoi cell from the barrier-free problem. The reason for this is that its perimeter is dictated by the barrier-influenced bisector $B^{\mathcal{B}}(P, Q) = \{X \in \mathbb{R}^2 \,|\, l_1^{\mathcal{B}}(P, X) = l_1^{\mathcal{B}}(Q, X)\}$ which, as we will see, can look very dissimilar to $B(P, Q)$. Whilst these new Voronoi cells and the overlying Voronoi diagram could be described with the prefix 'geodesic', in this paper we will refer to them simply as Voronoi cells $V(A_i)$ in the Voronoi diagram, and will specify otherwise if required.

Now we formally state the five location models chosen to be examined, each one focusing on finding the optimal location for the additional facility $A_{n+1} \in \mathcal{P}$.

In the Conditional Median Problem we seek the location A_{n+1} that minimises the total distance travelled

$$F_{CM}(A_{n+1}) = \int \int_{(u,v) \in \mathcal{P} \setminus int(\mathcal{B})} \min_{i=1,\ldots,n+1} l_1^{\mathcal{B}}(A_i, (u, v)) du dv, \tag{3.1}$$

cf. Cavalier and Sherali (1986). This is clearly appropriate for desirable facilities, though maximising the total distance travelled $F_{CM}(A_{n+1})$ is necessary when locating an undesirable facility (such as a rubbish dump). This is the Conditional Antimedian Problem. We note that it is clearly optimal to locate A_{n+1} at an existing facility so constrict the distance between the new and existing facilities to be no less than some threshold distance $D > 0$, as suggested by Berman and Huang (2008).

In the Conditional Centre Problem we seek the location A_{n+1} that minimises the maximum distance travelled by a customer

$$F_{CC}(A_{n+1}) = \max_{(u,v) \in \mathcal{P} \setminus int(\mathcal{B})} \min_{i=1,\ldots,n+1} l_1^{\mathcal{B}}(A_i, (u, v)). \tag{3.2}$$

This is important for facilities such as the emergency services.

In the Market Share Problem we seek the location A_{n+1} which, given $A_1, \ldots, A_k$ competitor facilities and $A_{k+1}, \ldots, A_n$ facilities of our own, maximises the total demand attracted by our facilities

$$F_{MS}(A_{n+1}) = \int \int_{(u,v) \in \{(u,v) \in \mathcal{P} \setminus int(\mathcal{B}) \,|\, \min_{i>k} l_1^{\mathcal{B}}((u,v), A_i) \leq \min_{i \leq k} l_1^{\mathcal{B}}((u,v), A_i)\}} 1 du dv, \tag{3.3}$$

cf. Eiselt, Pederzoli, and Sandblom (1985). This relies on the assumption that a customer will always have a preference for the nearest facility, choosing ours in the case of a tie.

In the Conditional Maximal Covering Location Problem we seek the location A_{n+1} which maximises the total demand captured of all customers no more than a certain distance away from a facility

$$F_{MCL}(A_{n+1}) = \int \int_{(u,v) \in \{(u,v) \in \mathcal{P} \setminus int(\mathcal{B}) \,|\, \min_{i=1,\ldots,n+1} l_1^{\mathcal{B}}((u,v), A_i) \leq R\}} 1 du dv \tag{3.4}$$

(cf. Matisziw and Murray (2009b)), where we assume that $R > 0$ is the maximum distance that a customer is willing to travel. For future ease define the unit ball of radius R about A_i to be $B_R(A_i)$.

4 Partitions of the Polygon that Preserve the Representation of the Objective Function

After adding the new generator A_{n+1} to the existing facilities $\{A_1, \ldots, A_n\}$, we call the new Voronoi diagram $\mathcal{VD}(A_1, \ldots, A_{n+1})$ the incremental Voronoi diagram. Some new edges and vertices that were not present in $\mathcal{VD}(A_1, \ldots, A_n)$ may appear in $\mathcal{VD}(A_1, \ldots, A_{n+1})$, and it is important that we understand how these changes affect the representation of the objective function as shown in the Introduction. If there is an area of the solution space $\mathcal{P}$ over which the representation of the objective function remains the same, then we are able to find extreme values over this area. To ease notation we will abbreviate $\mathcal{VD}(A_1, \ldots, A_n)$ as $\mathcal{VD}$.

In Averbakh et al. (2015), structural identity (concisely, the property of the Voronoi diagram having the same number of edges and vertices generated by the same points) was enough to guarantee the preservation of the representation. It was found that, after determining the Voronoi diagram for $\mathcal{P}$, three sets of lines induce a partition of $\mathcal{P}$ into $\mathcal{O}((n+m)^2)$ cells within which the representation of the objective function remains unchanged. Those lines are the configuration, intersection (the lines governing when and how a bisector intersects a bisector within $\mathcal{VD}$), and quadrant lines (the horizontal and vertical through the vertices of $\mathcal{P}$). For more detail see Averbakh et al. (2015). Therefore we have a polynomial number of cells over which we are able to find the optimal location A_{n+1} for a consistent demand function over the cells by exploiting first order methods.

However, these partitioning lines are not sufficient to ensure structural identity if a barrier is present. When a barrier is introduced in $\mathcal{P}$ we are almost guaranteed not to have representation identity within the partition provided by these lines. For a simple demonstration of the difference in representation within the partition obtained from Averbakh et al. (2015), consider one facility $A_1 = (0, 0)$ in a simple rectangular market region $\mathcal{P}$ within which a triangular barrier sits. Since $\mathcal{P}$ is a rectangle, the quadrant lines follow the perimeter of $\mathcal{P}$. For ease let us define

$$R(P, Q) = \left\{(x, y) \in \mathbb{R}^2 \mid \min[p_x, q_x] \leq x \leq \max[p_x, q_x],\ \min[p_y, q_y] \leq y \leq \max[p_y, q_y]\right\}$$

to be the rectangle with opposite corners $P = (p_x, p_y)$ and $Q = (q_x, q_y)$, and $\Delta(A, B, C)$ to be the triangle with vertices A, B, and C. Furthermore, since there is only one initial facility A_1 there are no bisectors within $\mathcal{VD}(A_1)$ to provide intersection lines. Therefore there are only configuration lines contributing to the partition. Figure 5 shows four placements of Z within the same partition cell (partitioning lines shown dashed) but with clearly different Voronoi diagram structures.

Thus we require more partitioning lines to be found in order to obtain a partition where the representation of the objective function is identical no matter where Z is located within each cell.

4.1 Barrier-Constrained Bisectors and Voronoi Diagrams

This motivates us to investigate constraints on the coordinates of Z such that the representation of the objective function is preserved. But before we can begin to look at such a thing we must properly understand how the barrier affects the Voronoi cells. To help classify the interaction between Z and $\mathcal{B}$, and to discuss these candidate partitioning cells for Z later, we introduce the following definitions.

Definition 4.1. *In a space with barriers, two points are considered visible from one another with respect to a norm if the shortest distance between the points is the same in the unconstrained and constrained space. The visible area of a space from a given point is the set of all points visible from that point, and the non-visible area is exactly its complement.*

Now that we have a notion of visibility, we will outline a particular area created by $\mathcal{B}$ which plays with this visibility.

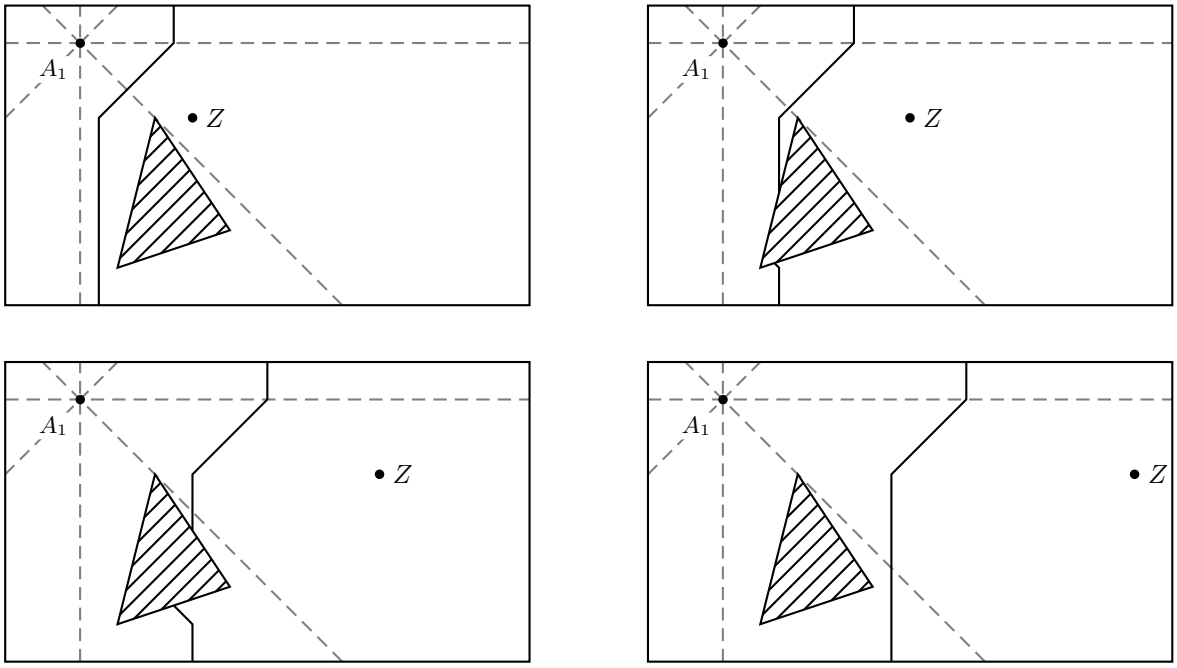
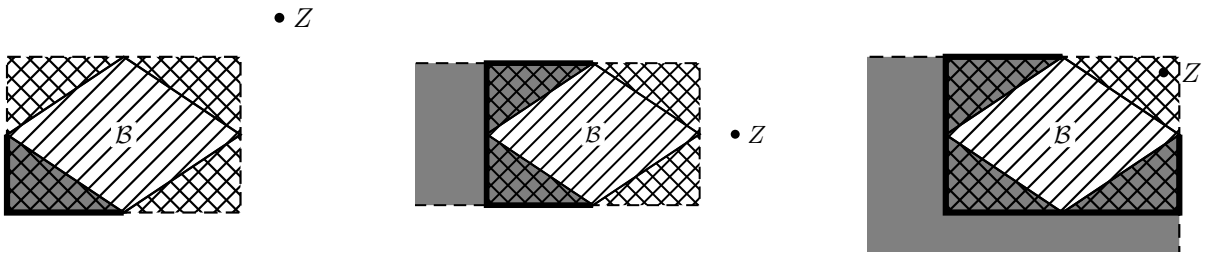


Figure 5: Four placements of Z within a rectangular $\mathcal{P} = R((-1, -3.5), (6, 0.5))$ with one existing facility $A_1 = (0, 0)$ and a triangular barrier $\mathcal{B} = \Delta((0.5, -3), (1, -1), (2, -2.5))$.

Definition 4.2. *For a p-sided polygonal barrier $\mathcal{B}$ with vertices $(b_{1x}, b_{1y}), \ldots, (b_{p_x}, b_{p_y})$, define its shadow in the l_1 norm to be*

$$Shadow(\mathcal{B}) = \left\{ (x, y) \in \mathbb{R}^2 : \min_i b_{ix} \leq x \leq \max_i b_{ix}, \min_i b_{iy} \leq y \leq \max_i b_{iy} \right\}$$

We call the leftmost, rightmost, topmost, and bottommost vertices of $\mathcal{B}$ the extreme points of $\mathcal{B}$ (and not just any vertex of $\mathcal{B}$ as in the classical definition) and the horizontal and vertical lines through these points the extreme lines; it is these which dictate the dimensions of $Shadow(\mathcal{B})$. Further, denote the corners of $Shadow(\mathcal{B})$ to be $C_1 = (\min_i b_{ix}, \max_i b_{iy})$, $C_2 = (\max_i b_{ix}, \max_i b_{iy})$, $C_3 = (\max_i b_{ix}, \min_i b_{iy})$, and $C_4 = (\min_i b_{ix}, \min_i b_{iy})$. We will find that the bisectors between two facilities interact not only with $\mathcal{B}$ but with $Shadow(\mathcal{B})$ and with the visible–non-visible boundary. We will use this term to describe the occurrence of certain bisectors, and it is important to understand the effect of this visibility as portrayed in Figure 6.



(a) Facility diagonal to $Shadow(\mathcal{B})$. (b) Facility adjacent to $Shadow(\mathcal{B})$. (c) Facility within $Shadow(\mathcal{B})$.

Figure 6: Visible areas of the space from different locations. Here a general barrier $\mathcal{B}$ is represented by a diamond. Grey areas are not visible from the facility.

To understand the effect that $\mathcal{B}$ has on the Voronoi diagram we need to know how it can change the bisector between two facilities. Now we introduce two important results regarding these bisectors in the barrier-constrained problem which we will make use of in our exploration. First of all we may

want to ask if, or when, the bisector is unchanged despite the addition of a barrier. Our first lemma sheds light, pardoning the pun, on this.

Lemma 4.1. *For a fixed trial location of the additional facility, any point on the bisector in the barrier-free facility location problem that lies in the visible areas as shown in Figure 6 for the existing and additional facility is a point on the bisector in the barrier-constrained facility location problem.*

Proof. In the barrier-free facility location problem, the boundary between the Voronoi cells of two facilities represents the set of points in the space that are equidistant from both facilities. If a point on this bisector lies in the visible area of a facility then it remains at the same distance from the facility as in the barrier-free problem. Therefore if a point on the bisector lies in both visible areas of the facilities then it is still equidistant from both facilities. Thus, it is a point on the bisector in the barrier-constrained facility location problem. □

Next we may wonder in how many instances a barrier can interact with one placement of Z, or rather how many times a bisector can intersect a barrier (though we will discover other interactions). The following lemma answers exactly this.

Lemma 4.2. *A bisector between any two facilities will intersect $\mathcal{B}$ twice or not at all (counting touches as not intersecting).*

Proof. Suppose the bisector $B^{\mathcal{B}}(P,Q)$ intersects $\mathcal{B}$ $n \geq 3$ times, splitting into several disconnected parts. If all these parts begin at $\mathcal{B}$ and extend to $\mathcal{P}$ then they partition the space into n regions, each region being a Voronoi cell of either facility P or Q. Each region will have exactly two neighbouring regions and be bounded by exactly one continuous boundary with each neighbour, a section of $\mathcal{B}$, and a section of $\mathcal{P}$. Since two Voronoi cells of the same facility must not be neighbours, n must be even. Therefore there must be at least two Voronoi cells of each facility, although this means that at least one Voronoi cell will not contain its respective facility (i.e. the shortest path to a point in that Voronoi cell of Q must travel through a Voronoi cell of P, and so contain points that are, by definition, closer to P than Q). This is a clear contradiction.

Alternatively, if not every disconnected part of the bisector extends to $\mathcal{P}$ then at least one part must intersect $\mathcal{B}$ at both ends. From an identical argument to that above, this region must contain a facility and, since it is surrounded by Voronoi cells of another facility, this becomes the only Voronoi cell, contradicting the supposition that the bisector intersects $\mathcal{B}$ more than twice.

Thus, the bisector may only intersect $\mathcal{B}$ twice or not at all. □

This lemma hints at one of the largest changes brought about by introducing a barrier. When the bisector intersects the barrier it splits into two different parts. Each of these parts corresponds to the shortest paths from at least one of the facilities circumnavigating either way around the barrier. In actual fact, these oriented shortest paths contribute to the bisector even when the bisector does not intersect the barrier. That is, in the presence of a barrier there is now a choice of paths to points in the non-visible area: one travels clockwise about the barrier, and the other anticlockwise; Lemma 4.1 showed that the length of the shortest paths between visible points remains unchanged, but with the introduction of a barrier we now have two path directions to consider between non-visible points. Let us define, for any P and Q non-visible or on the boundary of visibility from one another, $l_1^{\mathcal{B},+}(P,Q)$ and $l_1^{\mathcal{B},-}(P,Q)$ to be the shortest distances from P to Q in the geodesic l_1 norm clockwise or anticlockwise about $\mathcal{B}$ respectively. Now for any fixed point P and barrier $\mathcal{B}$, there will be areas of P's non-visible area for which the shortest paths from P to points within that area travel clockwise around $\mathcal{B}$ and other areas of the plane for which the shortest paths travel anticlockwise around $\mathcal{B}$. We call the set of points for which the shortest paths clockwise and anticlockwise around $\mathcal{B}$ are of equal length (i.e. the border of these two areas) the line of equidistance from P around $\mathcal{B}$; the line of equidistance from P around $\mathcal{B}$ is $\{X \mid l_1^{\mathcal{B},+}(P,X) = l_1^{\mathcal{B},-}(P,X)\}$. Figure 7 provides an example of a point, a barrier, and the types of possible shortest paths.

These two unique path orientations give rise to two potential contributions to the bisectors for $B^{\mathcal{B}}(P,Q)$: the points for which paths are equidistant travelling clockwise around $\mathcal{B}$ from P and anticlockwise around $\mathcal{B}$ from Q; and anticlockwise around $\mathcal{B}$ from P and clockwise around $\mathcal{B}$ from

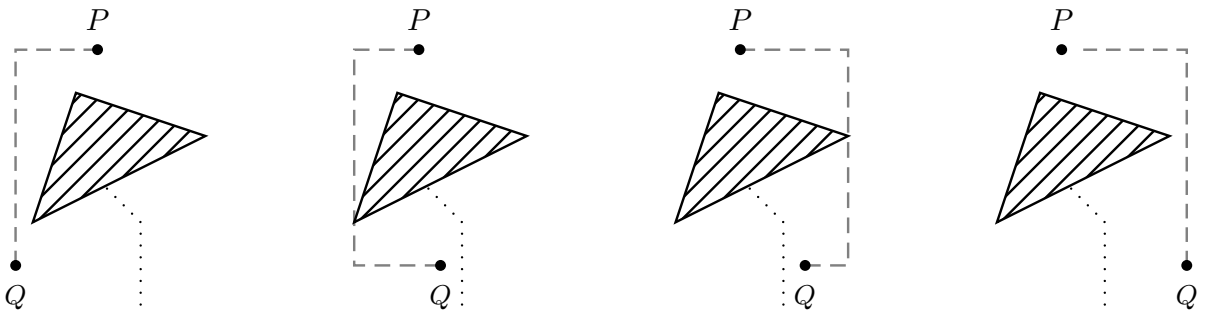


Figure 7: The four possible types of shortest paths between P and Q. The line of equidistance from P around $\mathcal{B}$ is shown as a dotted line.

Q. We shall refer to these potential contributions obtained by considering these oriented paths as non-visible bisector parts.

For any areas visible from P and non-visible from Q, we can obtain two unique non-visible bisector parts upon which the distance from P to a point on the bisector part is equal to the distance of the shortest path from Q to this point travelling clockwise or travelling anticlockwise about $\mathcal{B}$: $\{X \,|\, l_1(P,X) = l_1^{\mathcal{B},+}(Q,X)\}$ and $\{X \,|\, l_1(P,X) = l_1^{\mathcal{B},-}(Q,X)\}$ respectively. We can acquire analogous bisector parts for areas non-visible from P and visible from Q. In areas non-visible from both P and Q, we can seemingly again obtain two bisector parts $\{X \,|\, l_1^{\mathcal{B},+}(P,X) = l_1^{\mathcal{B},-}(Q,X)\}$ and $\{X \,|\, l_1^{\mathcal{B},-}(P,X) = l_1^{\mathcal{B},+}(Q,X)\}$. However, only one of these two bisector parts will be relevant. This is because, for the purposes of bisectors, for a fixed point X, we are only interested in the shortest paths to X from both of the facilities; unless X is on the line of equidistance around $\mathcal{B}$ from one of the facilities, these shortest paths will follow one orientation about $\mathcal{B}$. Figure 8 shows an example of the shortest paths in consideration, highlighting in Figure 8[left] how one pairing of oriented shortest paths is sensible while the other, in Figure 8[right], is nonsensical.

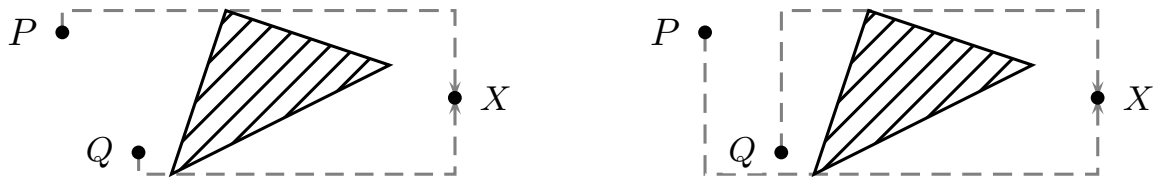


Figure 8: The shortest paths from P and Q to X travelling clockwise and anticlockwise around $\mathcal{B}$ respectively, and anticlockwise and clockwise around $\mathcal{B}$ respectively.

To discern which combination of directions is required for a general pair of facilities P and Q and barrier $\mathcal{B}$, we make use of the lines of equidistance from both P and Q about $\mathcal{B}$. On either side of these lines, the shortest paths from the facilities will travel clockwise, or anticlockwise, around $\mathcal{B}$. Thus the lines of equidistance can partition the space, non-visible from both facilities, into areas accessed clockwise from both facilities, accessed anticlockwise from both facilities, and accessed clockwise from one facility and anticlockwise from the other. It is only within this last area that bisector parts can exist so not only do the lines of equidistance tell us which orientations the shortest paths take, but they provide us with a feasible region within which we can find these bisector parts.

For an example, let us consider the area non-visible from P and Q (shaded purple) in Figure 9. The lines of equidistance show that this whole area is reached from P and Q by travelling clockwise and anticlockwise about $\mathcal{B}$ respectively in Figure 9[left]. Therefore we would produce the non-visible bisector parts $\{X \,|\, l_1^{\mathcal{B},+}(P,X) = l_1^{\mathcal{B},-}(Q,X)\}$. However, considering the placement of P and Q in Figure 9[middle], the line of equidistance from Q bisects the area non-visible from both facilities. The sub-area northwest of this line is reached via clockwise paths from both P and Q so we will not obtain

any bisector parts within this region (in actual fact, this area will clearly be contained within $V(P)$). Therefore we only consider $\{X \mid l_1^{\mathcal{B},+}(P,X) = l_1^{\mathcal{B},-}(Q,X)\}$ for X within the southeast portion of the purple region. If both lines of equidistance bisect the non-visible area, the region is partitioned further in Figure 9[right] and we would only consider the bisector parts $\{X \mid l_1^{\mathcal{B},+}(P,X) = l_1^{\mathcal{B},-}(Q,X)\}$ for X within the trapezium between the two lines of equidistance.

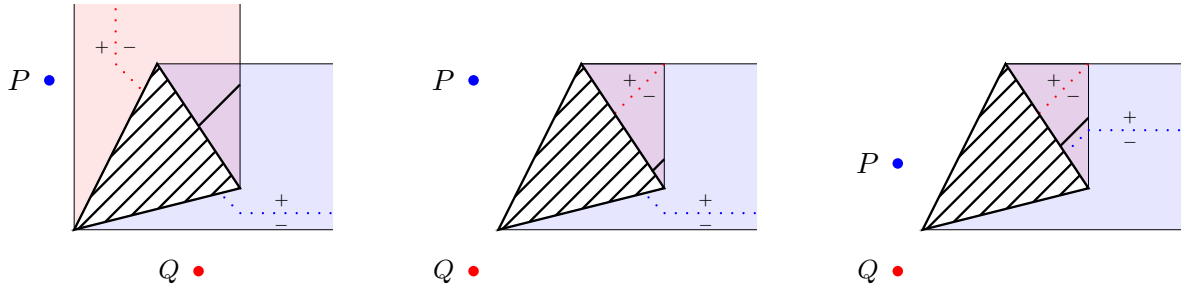


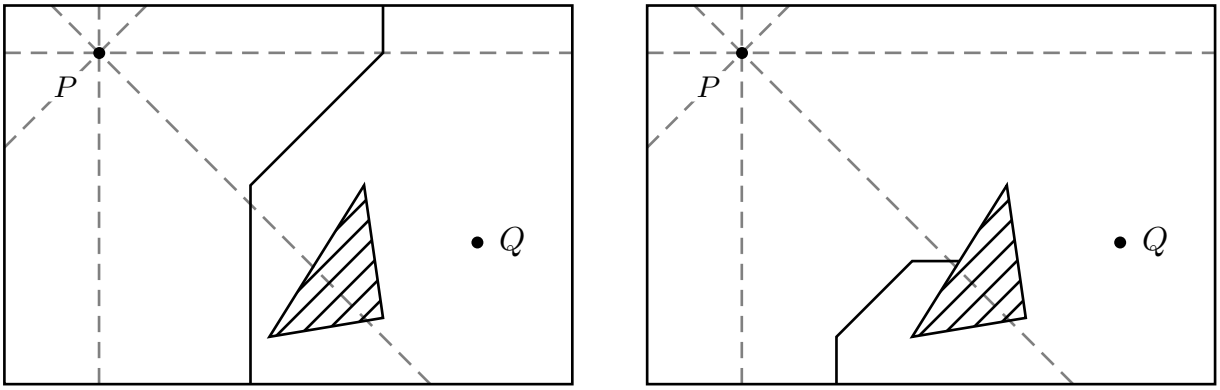
Figure 9: Non-visible bisector parts of P and Q. Non-visible areas from P and Q are shown shaded in their respective colours along with their line of equidistance around $\mathcal{B}$. Thus the area non-visible from both P and Q is shown in purple. The $+$ and $-$ annotate which side of the line of equidistance the shortest paths from the corresponding facility travel clockwise or anticlockwise about $\mathcal{B}$, respectively.

Now that we have detailed the construction of non-visible bisector parts, it is clear that these are instrumental in constructing geodesic bisectors since they correspond to points which are, according to particular paths, equidistant from P and Q, but how exactly is still a little unclear. Furthermore, it is uncertain how they connect to the region visible to both P and Q. To better explain this let us extend the non-visible bisector parts to the whole of $\mathcal{P}$. To do this, we note that the shortest paths that the bisector parts are defined by can be identified solely by their end points and the ordered extreme points of $\mathcal{B}$ which they must pass. Therefore, once a bisector part reaches the boundary of the visible area of (at least) one of the points P or Q, it will have met an extreme line through one of the extreme vertices of $\mathcal{B}$. If this extreme vertex is currently one defining the shortest path from the previously non-visible facility, the remainder of the bisector part within this visible area will not be defined by the passing of this extreme vertex. In fact, the shortest paths from this facility will be the original l_1 path (see above $\mathcal{B}$ in Figure 10a). On the other hand, if this extreme vertex is not defining the shortest path from the previous non-visible facility, the shortest paths from that facility will continue to pass the original extreme vertices. In this case the shortest path can be considered to be going 'the long way around' (see below $\mathcal{B}$ in Figure 10a). We refer to the result of the full continuation of the non-visible bisector parts beyond the boundary of the non-visible regions simply as bisector parts.

A complete example of these extended bisector parts is shown in Figure 10. Here we see the 'upper' bisector part in Figure 10a formed using the shortest path from Q anticlockwise about $\mathcal{B}$ and the 'lower' bisector part in Figure 10b from the clockwise path – note that these bisectors are piecewise linear and do not always partition the space.

This provides two bisector parts for any pair of points P and Q and barrier $\mathcal{B}$: that considering shortest paths going clockwise from P and/or anticlockwise from Q, and that of shortest paths anticlockwise from P and/or clockwise from Q. It is easy to discern the contribution of each part to the geodesic bisector when the bisector intersects $\mathcal{B}$ (identify, for example, the separation of the two bisector parts in Figure 5). However, these bisector parts still exist when no intersection occurs. In fact, the geodesic bisector is exactly the combination of each bisector part contained within its respective side of the appropriate line of equidistance, since crossing this line of equidistance means that the opposite oriented path is preferable, and so the other bisector part is used. It is always along this line that the bisector swaps from the 'upper' part to the 'lower' part. The bisector parts from Figure 10 are combined in this way in Figure 11 to create the full bisector $B^{\mathcal{B}}(P,Q)$.

Therefore to understand the bisector we must understand the bisector parts; the bisector's representation depends entirely on the representation of the bisector parts and on which of their segments



(a) 'Upper' bisector part: $l_1(P,X) = l_1^{\mathcal{B},-}(Q,X)$.

(b) 'Lower' bisector part: $l_1(P,X) = l_1^{\mathcal{B},+}(Q,X)$.

Figure 10: The different bisector parts for the placement of $Q = (4,-2)$ within $\mathcal{P} = R((-1,-3.5),(5,0.5))$ with one existing facility $P = (0,0)$ and $\mathcal{B} = \Delta((1.8,-3),(2.8,-1.4),(3,-2.8))$.

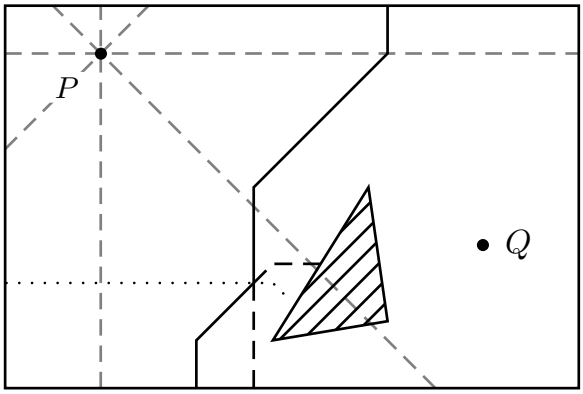


Figure 11: The combination of the bisector parts to create $B^{\mathcal{B}}(P,Q)$ for the placement of $Q = (4,-2)$ within $\mathcal{P} = R((-1,-3.5),(5,0.5))$ with one existing facility $P = (0,0)$ and $\mathcal{B} = \Delta((1.8,-3),(2.8,-1.4),(3,-2.8))$. The line of equidistance from Q about $\mathcal{B}$ is shown as a dotted line.

the line of equidistance around $\mathcal{B}$ intersects. We have formally defined the bisector parts, but how exactly do they behave? For this we must explore what a bisector part resembles within a non-visible area. As discussed previously in our extension of bisector parts to $\mathcal{P}$, for a bisector part within a non-visible area the shortest path from one of the facilities must always pass by an extreme point of $\mathcal{B}$ (see Figure 6). This extreme point then acts as one of the facilities for the bisector part beyond this point, but rather than the bisector part being the set of points that are equidistant from the visible facility and the extreme point (according to the certain shortest path orientation about $\mathcal{B}$), it is the set of points that are closer to the extreme point than to the visible facility by a magnitude of exactly twice the distance between the extreme point and the non-visible facility. These are called additive bisectors (Okabe et al., 2000).

More rigorously, in the barrier-free case the additive bisector between P and Q with additive constant c is $B_c^{add}(P,Q) = \{X \in \mathbb{R}^2 \,|\, l_1(P,X) = l_1(Q,X) + c\}$. These have six unique forms as shown in Figure 12, with the sixth being the empty cell if $|c| > l_1(P,Q)$. As we can see, a value of $c = 0$ gives the original bisector.

It is worth addressing the fact that two of these additive bisectors produce quarter planes. This will only happen if $|c| = ||p_x - q_x| - |p_y - q_y||$ which is rare, but of course not impossible. We discuss how we can deal with the quarter planes at the end of this section.

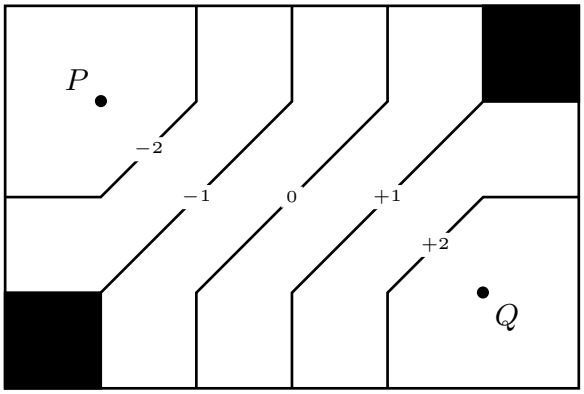


Figure 12: Additive bisectors $B_c^{add}(P, Q)$ for various values of c (as labelled) where $P = (0, 0)$ and $Q = (2, -1)$.

Thus, when the bisector part between P and Q enters a non-visible area from Q (i.e. the shortest paths must all pass by an extreme point B_i of $\mathcal{B}$) the bisector part in this non-visible area will be the additive bisector between P and B_i with c equal to the distance the shortest path has travelled from Q to B_i. In the barrier-constrained setting taking, for example, the bisector part from the clockwise path from P and anticlockwise from Q, we have the geodesic additive bisector $B_c^{add,+}(P, B_i) = \{X \in \mathbb{R}^2 \,|\, l_1^{\mathcal{B},+}(P, X) = l_1^{\mathcal{B},-}(B_i, X) + c\}$ where $c = l_1^{\mathcal{B},-}(B_i, Q)$. The analogous definition for the 'lower' part $B_c^{add,-}(P, B_i)$ is clear from this. Using Figure 10b for a numerical example, the bottommost vertex of $\mathcal{B}$, $B = (1.8, -3)$, acts as the new active point within the additive bisector. Here $c = l_1(Q, B) = 3.2$ so Figure 10b displays $B_{3.2}^{add,-}(P, B) = \{X \in \mathbb{R}^2 \,|\, l_1^{\mathcal{B},-}((0, 0), X) = l_1^{\mathcal{B},+}((1.8, -3), X) + 3.2\}$.

This gives us a closer idea of how the bisector part looks when it enters a non-visible area from either of the original facilities. This can obviously be extended to how it looks whenever it enters any new non-visible areas from any of the active points of the bisector part (facilities or extreme points from which the additive bisector is acting) as it will simply create a new additive bisector. Thus the bisector parts are a concatenation of the original bisector and additive bisectors becoming a new additive bisector whenever the current one enters an area not visible from either one of the active points. An example of these bisector parts contributing to the final bisector is shown in Figure 13 (observe that an additive bisector may even have its own additive bisector).

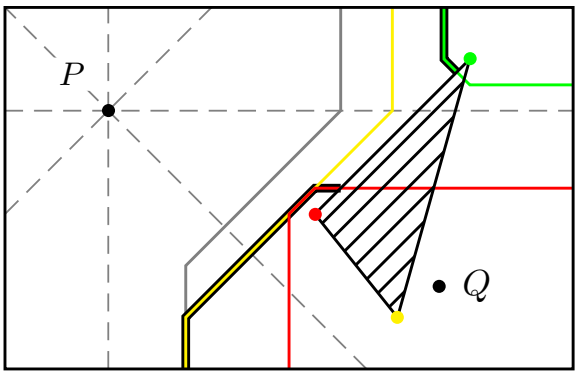


Figure 13: Bisectors and additive bisectors for the facilities $P = (0, 0)$ and $Q = (3.2, -1.7)$ in the presence of $\mathcal{B} = \Delta((2, -1), (2.8, -2), (3.5, 0.5))$. Each additive bisector $B_c^{add}(P, B_i)$, for B_i a vertex of $\mathcal{B}$ and $c = l_1^{\mathcal{B}}(B_i, Q)$, is colour-coded to match B_i. Moreover, the bisector in the original barrier-free problem is shown in grey, and the overall barrier-constrained bisector is highlighted with a black border.

It is important to note that the breakpoints (and therefore the edges) of additive bisectors are linear in the coordinates of original facilities: Q contributes linearly to c ($= l_1^{\mathcal{B},+}(B_i, Q)$ or $= l_1^{\mathcal{B},-}(B_i, Q)$) which has a linear effect on the additive bisector's breakpoints. So we have defined all potential Voronoi cells through the bisector parts of their bisectors. The exact production of the specific bisectors is studied in more detail in Appendix A.4, and the configuration lines used to determine the expression of the additive bisector (as shown in Figure 12) are derived in Section 4.2.4.

Lemma 4.3. *A bisector's direction will only change by* $45°$ *unless it crosses the line of equidistance around* $\mathcal{B}$ *in which case it can also change by* $90°$.

Proof. See Appendix A.1. □

It is worth noting that even though the bisector parts may be affected by the barrier their combination to form the bisector may not be.

Now it may be sensible to ask how the properties of the Voronoi diagram in the barrier-constrained problem differ from that in the barrier-free problem. The properties of interest are the shapes of the cells and the number of edges and vertices in the Voronoi diagram. Through our foray into barrier-constrained bisectors (which contribute to the edges of the Voronoi diagram) we have seen that Voronoi cells are no longer star-shaped. However, they do remain connected: for any $P \in \mathcal{P}$ and facility A_i, if the shortest path between A_i and P passes through $P' \in int(V(A_j))$, say, then

$$l_1^{\mathcal{B}}(A_i, P) = l_1^{\mathcal{B}}(A_i, P') + l_1^{\mathcal{B}}(P', P) > l_1^{\mathcal{B}}(A_j, P') + l_1^{\mathcal{B}}(P', P) \geq l_1^{\mathcal{B}}(A_j, P)$$

so $P \notin V(A_i)$. Further more detailed exploration of the bisectors can be found later in Section 4.2. We shall concern ourselves now with the number of vertices and edges of barrier-constrained Voronoi diagrams, as answered in the following lemma.

Lemma 4.4. *For facilities* $A_1, \ldots, A_n$ *and a* p*-sided convex barrier* $\mathcal{B}$ *located with an* m*-sided convex polygon* $\mathcal{P}$*, the number of edges and the number of vertices in* $\mathcal{VD}(A_1, \ldots, A_n)$ *are each* $\mathcal{O}(m+n+p)$.

Proof. See Appendix A.2. □

For our facility location problems, the representation of the objective function changes if the representation of the Voronoi diagram changes. Therefore, using the knowledge obtained above, we explore the different structures that the Voronoi diagram can take, and look for a way to categorise these and identify the partition lines across which these changes take place.

Firstly we experiment with how a simple triangular barrier can divide up the space into regions that, if Z is placed within the region, share the same representation of the bisector. An example of this is shown in Figure 15b where every cell corresponds to a different Voronoi diagram representation (the solid grey space covers the area over which the Voronoi diagram is not affected by the barrier). From sketches such as these one can see that the effect of the barrier depends entirely on the location of Z with the bisector remaining unaffected if Z is located in a large area of $\mathcal{P}$. There are numerous approaches we could take to describe partitions like this for a general $\mathcal{B}$. Equations were found for the partition lines depending on all of Z's possible positions in relation to A_1 and $\mathcal{B}$; however, one can rightly imagine that this would involve many separate cases even when restricted to the $1+1$ case. The simplicity and suggested patterns of diamond segments exhibited in Figure 15b hint at an easier solution. Therefore we are interested in how we can define the perimeter lines for each partition cell (specifically, what it means for Z to be located upon one of the lines) so that we can easily construct such a partition for any barrier $\mathcal{B}$.

Now the stage is set for us to introduce the categories of partition lines and their general forms.

4.2 Conditions Preserving the Representation of the Objective Function

Having investigated what it means for a facility to lie on a partitioning line, for any convex $\mathcal{P}$ with vertices $P_i = (p_{ix}, p_{iy})$ for $i = 1, \ldots, m$ and any convex $\mathcal{B}$ with vertices $B_i = (b_{ix}, b_{iy})$ for $i = 1, \ldots, p$ we can break the lines down into seven cases:

- configuration lines of A_k;
- additive configuration lines of A_k;
- extreme lines of $\mathcal{B}$;
- geodesic diamonds centred on vertices of $\mathcal{VD}$;
- intersection lines of breakpoints and edges of $\mathcal{VD}$;
- kink lines of the corners of $Shadow(\mathcal{B})$;
- wrap-around lines of $\mathcal{B}$.

We will detail each of these lines separately, deriving general expressions for the lines with n existing facilities (without loss of generality we take $A_1 = (0, 0)$ for ease of calculation unless stated otherwise), and prove that they are necessary and sufficient for forming the full partition.

4.2.1 Configuration lines

It was shown in Averbakh et al. (2015) that these lines, stated below for $A_1 = (0, 0)$ for completeness:

$$y = 0\,,\ \ x = 0\,,\ \ y = x\,,\ \ y = -x\,,$$

were necessary for a full partition. Their use has been demonstrated earlier in Section 3, and illustrated in Figure 4. While we could include all of the lines identified in that paper for our partition (since the barrier-free problem can be seen as a problem contained in the barrier-constrained problems – for a barrier of negligible size, or one outside $\mathcal{P}$, say), equivalent such lines for all barrier-constrained problems have been found in this work. Therefore, stating the other lines found in Averbakh et al. (2015) would be unnecessary and simply complicate the partition.

While, following on from the configuration lines for bisectors unobstructed by $\mathcal{B}$, this may seem a natural place to define additive configuration lines, the theory of the following lines proves useful in the derivation of additive configuration lines so we will return to these in Section 4.2.4.

4.2.2 Extreme lines

Our first new lines are the easiest to define and have already been introduced. The extreme lines of $\mathcal{B}$ are simply

$$x = \min b_{ix}\,,\ \ x = \max b_{ix}\,,$$
$$y = \min b_{iy}\,,\ \ y = \max b_{iy}\,.$$

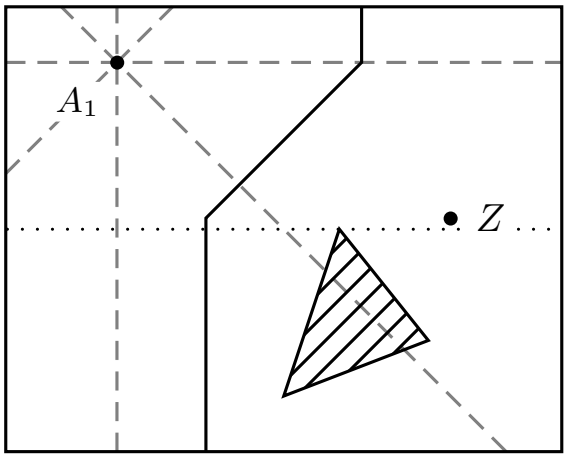
These lines delineate where the shortest path from Z to certain areas of $\mathcal{P}$ may now be restricted to pass by a vertex of $\mathcal{B}$ (since $\mathcal{B}$ is convex a shortest path will only have to pass a vertex of $\mathcal{B}$ if that vertex is an extreme one, and from an extreme point all points are visible till the next extreme point restricts it). Since the representation of the shortest paths changes, so can the bisector. We will show that these lines are necessary by way of an example shown in Figure 14, where the lower bisector breakpoint is on the same horizontal line as Z in Figure 14a, but not in Figure 14b.

For any barrier $\mathcal{B}$ this gives exactly four lines.

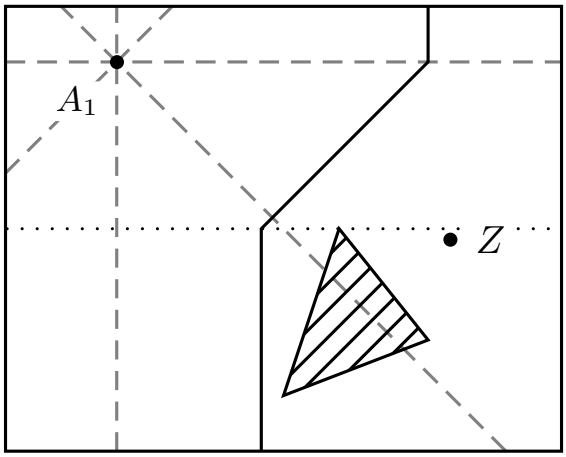
4.2.3 Geodesic diamonds

Next we will introduce geodesic diamonds, and explore them about vertices B_i of $\mathcal{B}$ – though the results for other vertices within $\mathcal{VD}$ follow identically.

While working with the l_1 norm, the gauge (lines of equidistance) is the diamond centred about its point of symmetry. Looking at existing partitions these diamonds now become blindingly obvious – however, importantly, not all are diamonds and some appear to have been distorted. This distortion is due to the barriers which, as we know all too well, can interrupt the shortest paths. In this way



(a) Bisector with coordinates $(\frac{x+y}{2}, -3.5)$, $(\frac{x+y}{2}, y)$, $(\frac{x-y}{2}, 0)$, and $(\frac{x-y}{2}, 1)$.



(b) Bisector with coordinates $(\frac{x-y-3}{2}, -3.5)$, $(\frac{x-y-3}{2}, -1.5)$, $(\frac{x-y}{2}, 0)$, and $(\frac{x-y}{2}, 1)$.

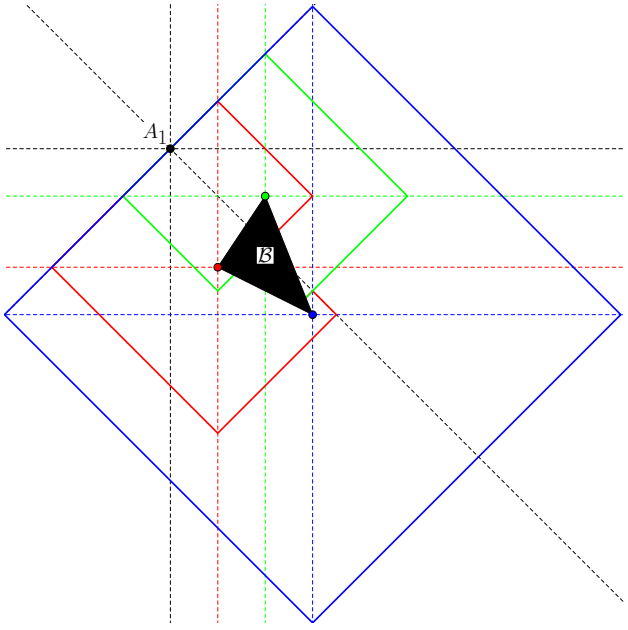
Figure 14: Two placements of $Z = (x, y)$ within $\mathcal{P} = R((-1, -3.5), (4, 0.5))$ with one existing facility $A_1 = (0, 0)$ and $\mathcal{B} = \Delta((1.5, -3), (2, -1.5), (2.8, -2.5))$.

these partitioning lines are the geodesic gauges with their centre at a vertex of the barrier, and size given by the distance from that vertex to A_1.

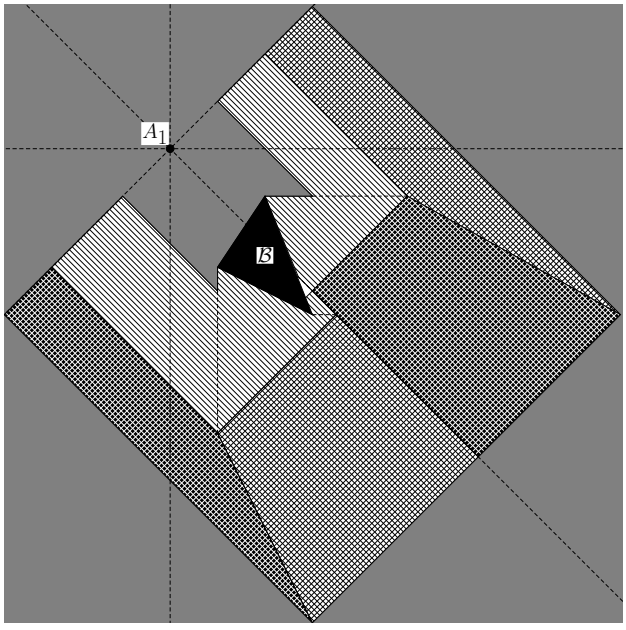
We define the geodesic diamond centred on X to A_1 (as in, A_1 lies on the geodesic diamond) to be

$$\diamond_g^X(A_1) := \{Q \in \mathcal{P} \setminus \mathcal{B} : l_1^{\mathcal{B}}(Q, X) = l_1^{\mathcal{B}}(A_1, X)\}$$

where, here, we will take X to be one of the vertices of $\mathcal{VD}$.



(a) Geodesic diamonds centred on each barrier vertex to A_1.



(b) Complete partition wherein each cell has structurally independent Voronoi diagrams.

Figure 15: The geodesic diamonds and complete partition of a large $\mathcal{P}$ with one existing facility $A_1 = (0, 0)$ and $\mathcal{B} = \Delta((2, -5), (4, -2), (6, -7))$.

It is on these diamonds that B_i is equidistant from A_1 as it is from Z and so, by definition, will belong on the bisector. Therefore they delineate where the bisector will intersect a new line (edge of

$\mathcal{VD}$) and thus lead to a change in the cell structure. For any vertex of $\mathcal{VD}$, it is important to note that we need only concern ourselves with the geodesic diamond to the facility generating the Voronoi cell within which the vertex lies. So each vertex of $\mathcal{VD}$ contributes exactly one geodesic diamond. From Figure 15 we can see that these geodesic diamonds contribute almost in their entirety to the majority of that partition. Indeed these geodesic l_1 gauges can be more colourful, as we see that they change direction by 90° whenever they hit a quadrant line (a vertical or horizontal) of its centre vertex or an extreme line of $\mathcal{B}$.

We now include a lemma to count the maximum number of lines.

Lemma 4.5. *For a convex $\mathcal{B}$, geodesic diamonds have a maximum of seven edges.*

Proof. See Appendix A.3. □

In Figure 16 we illustrate the possible cases for geodesic diamonds centred upon a vertex of $\mathcal{B}$ either at a corner of $Shadow(\mathcal{B})$, on an edge of $Shadow(\mathcal{B})$, or within the interior of $Shadow(\mathcal{B})$.

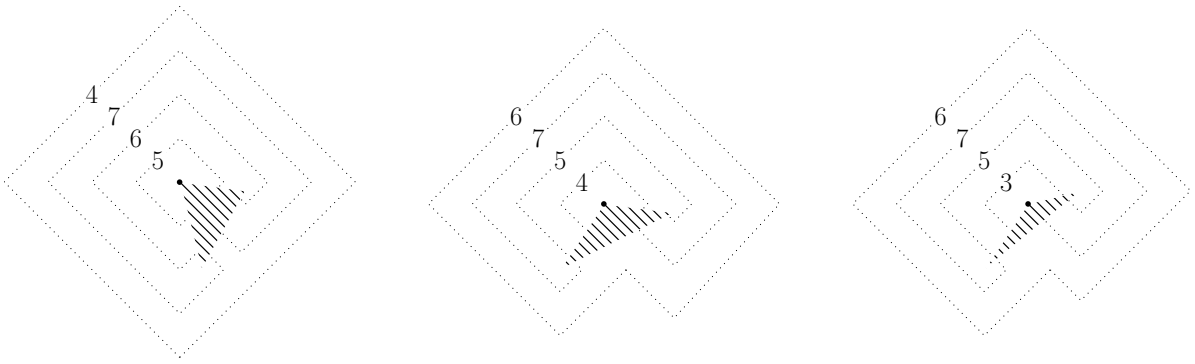


Figure 16: The possible geodesic diamonds centred at corner, edge, or interior vertices of $\mathcal{B}$.

Thus we obtain at most $\mathcal{O}(m+n+p)$ lines of this type: each $\mathcal{O}(m+n+p)$ vertex of $\mathcal{VD}$ contributes a geodesic diamond of at most seven edges.

4.2.4 Additive configuration lines

Just as the configuration lines are required to determine what orientation the bisector $B(A_i, Z)$ takes, we must find the additive configuration lines which dictate the type of additive bisector part present between facilities A_i and Z, as identified in Figure 12.

We wish to identify the lines across which the movement of an additional facility Z would change the type of additive bisector present (as portrayed in Figure 12). Fixing one facility, these lines correspond to the placement of the free facility so that the additive bisector creates a quarter plane (as this differentiates between the appropriate non-degenerate additive bisectors). Equivalently, for $B_c^{add}(P, Q)$, we find the points at which the non-active point corners of $R(P, Q)$ are equidistant from A_i and Z. It is important to note that the structure of the additive bisector can change in the way described only when we consider the additive bisector between active points which are visible from one another. We will split the additive bisectors possible in our set-up into three cases.

Firstly we examine additive bisectors between facilities A_i and Z whose active points B_j and B_k (respectively) are vertices of $\mathcal{B}$: $B^{add}_{l_1^{\mathcal{B}}(B_k,Z)}(B_j, B_k)$. For this we require B_k to be non-visible from A_i and B_j to be non-visible from Z (else one of the facilities would be an active point). Since B_j and B_k are visible from one another and are extreme points of $\mathcal{B}$, $R(B_j, B_k)$ must share at least one corner $C_l \notin \{B_j, B_k\}$ with $Shadow(\mathcal{B})$. We must have $C_l \notin \mathcal{B}$ else there would be no such active points B_j or B_k. This C_l is unique unless $R(B_j, B_k) = Shadow(\mathcal{B})$, in which case we do not choose the corner in the interior of the visible region from A_i and Z. By the convexity of $\mathcal{B}$, this must be the only corner from which a quarter plane can begin (the shortest path from A_i or Z to the alternative corner – if it exists, i.e. the corner may lie within $\mathcal{B}$ – does not pass B_j or B_k respectively so cannot exist in $B^{add}_{l_1^{\mathcal{B}}(B_k,Z)}(B_j, B_k)$). Therefore the additive bisector creates a quarter plane if C_l is equidistant from A_i

and Z. For a general A_i and $\mathcal{B}$ we require C_l to be either non-visible from A_i or on the boundary of A_i's visibility (for B_k to be non-visible from A_i). For $A_i \in Shadow(\mathcal{B})$ there can be as many as three such C_l, otherwise there are a maximum of two.

Thus, for any facility A_i, for a corner C_l either non-visible or on the boundary of the visible area from A_i we produce the lines

$$l_1^{\mathcal{B}}((x,y),C_l) = l_1^{\mathcal{B}}(A_i,C_l)\,.$$

This is a geodesic diamond $\diamond_g^{C_l}(A_i)$; however, we only require the segments contained within the quadrant of C_l containing $\mathcal{B}$ (since C_l must be non-visible, or on the boundary of the visible region, from Z as well as A_i). Moreover, we only consider the segments of $\diamond_g^{C_l}(A_i)$ on the opposite side to A_i of the line of equidistance from C_l around $\mathcal{B}$ (since the shortest path to C_l from A_i and Z must travel in opposite directions around $\mathcal{B}$ for B_j to be closer to A_i than to Z and for B_k to be closer to Z than to A_i). The line of equidistance from C_l around $\mathcal{B}$ is always a quarterplane at the opposite corner of $Shadow(\mathcal{B})$ to C_l (being the quadrant opposite to the one containing $\mathcal{B}$) and the diagonal line from this corner towards $\mathcal{B}$. This amounts to a maximum of two line segments (as can be seen from the geodesic diamonds centred on a corner vertex – i.e. C_l – in Figure 16). An example of such a case and its partitioning line is shown in Figure 17.

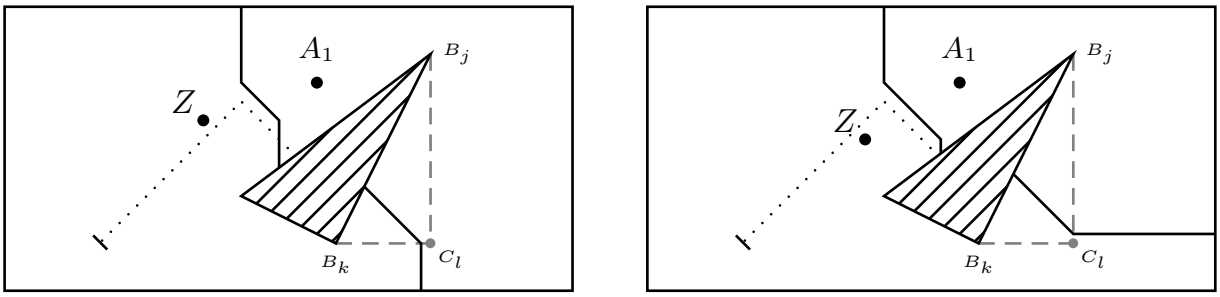


Figure 17: Two placements of Z within $\mathcal{P} = R((-1.5,0),(4.5,3))$ along with facility $A_1 = (1.8,2.2)$ and $\mathcal{B} = \Delta((1,1),(2,0.5),(3,2.5))$.

Secondly we will examine $B^{add}_{l_1^{\mathcal{B}}(B_j,Z)}(A_i,B_j)$, the barrier-constrained bisector between facilities A_i and Z where the additive bisector part acts between $A_i = (a_x,a_y)$ and the vertex $B_j = (b_x,b_y)$ of $\mathcal{B}$. For this we require B_j to be an extreme vertex of $\mathcal{B}$ visible from A_i (otherwise it would be a bisector of the previous type). Note that B_j can either lie on the perimeter of the non-visible area from Z or be non-visible from Z (being the second, or even third, active point which must be passed in the shortest path). Now since the active point A_i of this additive bisector is a facility, we will always have $c > 0$. This means that the additive bisector will always have a breakpoint on the longer edge of $R(A_i,B_j)$ adjacent to B_j and it remains to be shown whether the other breakpoint lies on the other longer edge or the other edge adjacent to B_j.

Therefore the corner of interest of $R(A_i,B_j)$ is the one closest to B_j, and this depends on the configuration cone of A_i that B_j lies within. If $B_j \in \mathcal{CC}^1(A_i) \cup \mathcal{CC}^4(A_i) \cup \mathcal{CC}^5(A_i) \cup \mathcal{CC}^8(A_i)$ then the corner of interest is (b_x,a_y), otherwise it is (a_x,b_y). If this corner is not visible from either of the active points due to $\mathcal{B}$ obscuring access then this additive bisector is not able to produce differing types of bisectors (we need not concern ourselves with $\mathcal{P}$ as, whilst it may cut out the corner of interest, it does not affect the shortest paths). In this case we must explore the additive bisector between the new active points. The same is, or can be, true if the incident edge of $\mathcal{B}$ follows the perimeter of $Shadow(\mathcal{B})$. Thus, the corner must be visible from both A_i and B_j for the configuration to be able to change. Therefore additive bisector changes of this type can only occur for B_j, an extreme vertex lying on an extreme line of $\mathcal{B}$ which would separate $\mathcal{B}$ and A_i. For any A_i and $\mathcal{B}$ there can be at most two vertices of this type (since there are only a maximum of two extreme lines of $Shadow(\mathcal{B})$ which could separate A_i from $\mathcal{B}$ and an edge of $\mathcal{B}$ must not follow the outline of $Shadow(\mathcal{B})$ as described above).

Now, as before, for this corner to be on the bisector, $Z = (x, y)$ must satisfy

$$l_1^{\mathcal{B}}((x,y),(b_x,a_y)) = l_1^{\mathcal{B}}(A_i,(b_x,a_y)) = |b_x - a_x| \text{ or } l_1^{\mathcal{B}}((x,y),(a_x,b_y)) = l_1^{\mathcal{B}}(A_i,(a_x,b_y)) = |b_y - a_y|$$

depending on the location of B_j with respect to A_i as discussed above. Once more this is a geodesic diamond $\diamond_g^{(b_x,a_y)}(A_i)$ (or $\diamond_g^{(a_x,b_y)}(A_i)$) but it is composed of only the points for which the shortest paths from (b_x, a_y) (or (a_x, b_y)) must travel past B_j. That is, it is restricted by the extreme line $x = b_x$ or $y = b_y$ (respectively), the line of equidistance from (b_x, a_y) or (a_x, b_y) (respectively) around $\mathcal{B}$, and $\mathcal{B}$ itself. Now this shortest path may pass as many as three extreme vertices before it reaches the line of equidistance or $\mathcal{B}$ (passing four would cause it to enter an area visible from the starting point so it must have crossed the line of equidistance). Thus it consists of a maximum of three line segments. The effect of crossing this line is shown in Figure 18.

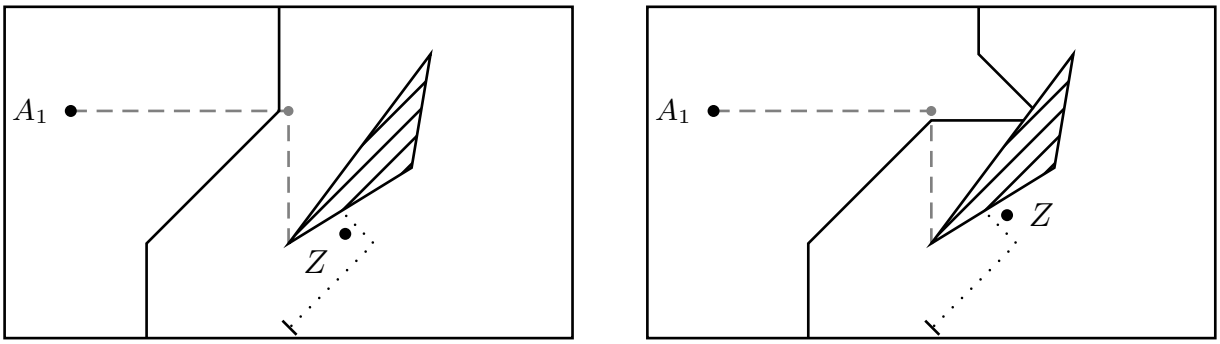


Figure 18: Two placements of Z within $\mathcal{P} = R((-1.5, 0), (4.5, 3.5))$ along with facility $A_1 = (-0.8, 2.4)$ and $\mathcal{B} = \Delta((1.5, 1), (2.8, 1.8), (3, 3))$.

Finally, we will examine $B^{add}_{l_1^{\mathcal{B}}(A_i,B_j)}(Z, B_j)$, the barrier-constrained bisector between facilities A_i and Z where the additive bisector part acts between $Z = (x, y)$ and the vertex $B_j = (b_x, b_y)$ of $\mathcal{B}$. For this we require B_j to be an extreme vertex of $\mathcal{B}$ within or on the boundary of the non-visible area from A_i, or the previous active point (otherwise B_j is not an active point of the additive bisector). Moreover we only consider Z visible from B_j and not from A_i. As before, since the active point Z of this additive bisector is a facility, we will always have $c > 0$. So we are interested in the locations of Z for which the bisector intersects the corner of $R(Z, B_j)$ closest to B_j. But since Z can move, this corner may be either the one lying on the horizontal or the vertical through B_j depending on Z's location.

Therefore, for every half-line travelling horizontally or vertically from B_j, within or on the boundary of the non-visible area from A_i or the previous active point, we must consider Z being located such that the closest corner of $R(Z, B_j)$ to B_j lies on this line and the additive bisector travels through it. That is, for $Z = (x, y)$ lying in this area,

$$l_1^{\mathcal{B}}((x,y),(b_x,y)) = l_1^{\mathcal{B}}(A_i,(b_x,y)) \Rightarrow |y - b_y| = |x - b_x| - l_1^{\mathcal{B}}(A_i, B_j) \text{ or}$$
$$l_1^{\mathcal{B}}((x,y),(x,b_y)) = l_1^{\mathcal{B}}(A_i,(x,b_y)) \Rightarrow |y - b_y| = |x - b_x| + l_1^{\mathcal{B}}(A_i, B_j)$$

for the horizontal and vertical cases discussed respectively. While this gives the equation of eight lines, each one must lie in the non-visible area from A_i (or the previous active point if existing) and begin at the extreme lines through B_j. Thus only half of these at most will satisfy these conditions. Furthermore, each of these lines is only exhibited if the configuration cone $\mathcal{CC}^k(B_j)$ containing the line has $\left(\mathcal{CL}^k(B_j) \cup \mathcal{CL}^{k+1}(B_j)\right) \cap \mathcal{B} = \{B_j\}$ (else the bisectors will present the same since the bisector will intersect $\mathcal{B}$ before changing configuration). So, by the convexity of $\mathcal{B}$, this means we require a maximum of only three of these partition lines. The effect of crossing these lines is shown in Figure 19 in the lower bisector part of $B^{\mathcal{B}}(Z, A_1)$.

Since these lines are important to categorise, we shall state them as a result as follows.

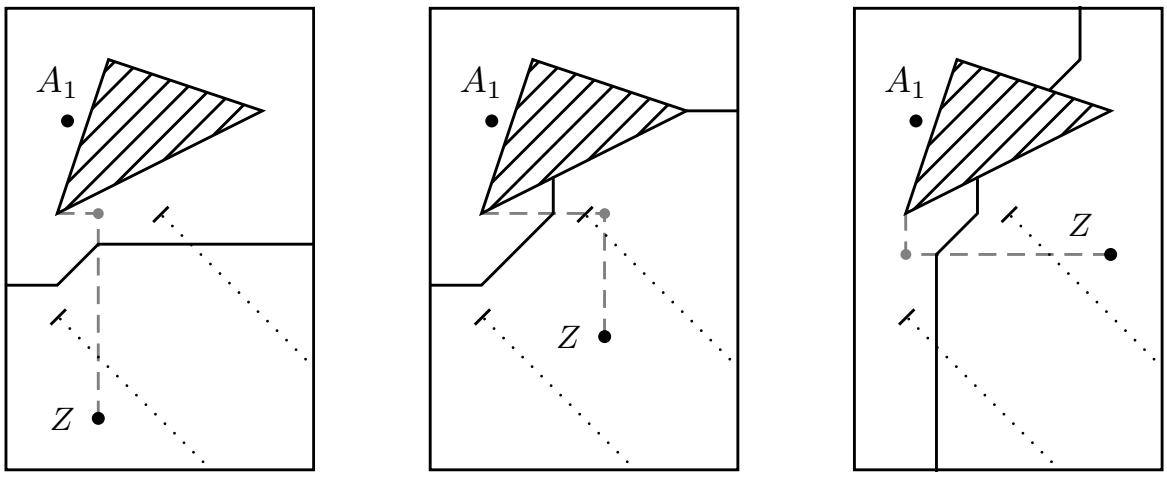


Figure 19: Three placements of Z within $\mathcal{P} = R((0.5, -0.5), (3.5, 4))$ along with facility $A_1 = (1.1, 2.9)$ and $\mathcal{B} = \Delta((1, 2), (1.5, 3.5), (3, 3))$.

Lemma 4.6. *For a fixed facility $A_i = (a_x, a_y)$ and barrier $\mathcal{B}$, the identification of the additive bisectors between A_i and $Z = (x, y)$ is determined by the following additive configuration lines:*

- $\diamond_g^{C_l}(A_i)$ *for $C_l \notin \mathcal{B}$ non-visible or on the boundary of the visible area from A_i,*
- $\diamond_g^{(b_x,a_y)}(A_i)$ *or* $\diamond_g^{(a_x,b_y)}(A_i)$ *for $B_j = (b_x, b_y)$ on an extreme line of $\mathcal{B}$ that separates A_i from $\mathcal{B}$, if $B_j \in \mathcal{CC}^1(A_i) \cup \mathcal{CC}^4(A_i) \cup \mathcal{CC}^5(A_i) \cup \mathcal{CC}^8(A_i)$ or otherwise (respectively),*
- $|y - b_y| = |x - b_x| \pm l_1^{\mathcal{B}}(A_i, B_j)$ *for any extreme vertex $B_j = (b_x, b_y)$ of $\mathcal{B}$.*

These additive configuration lines contribute to an overall count of $\mathcal{O}(n)$ line segments:

- a maximum of two line segments from every line of the first case, for which a facility A_i has a maximum of three possible corners C_k with which to produce a line (so a maximum of $6n$ possible lines of the first case),
- a maximum of three line segments from every line of the second case, for which a facility A_i has a maximum of two possible extreme vertices B_j with which to produce a line (so a maximum of $6n$ possible lines of the second case), and
- a maximum of three half-lines of the third case, for which a facility A_i has a maximum of four possible extreme vertices B_j with which to produce half-lines (so a maximum of $12n$ possible lines of the third case).

4.2.5 Breakpoint intersection lines

The next category of lines occurs when a breakpoint of the bisector intersects an edge of $\mathcal{VD}$. We will begin by demonstrating breakpoint intersection lines for when a breakpoint intersects an edge of $\mathcal{B}$.

A breakpoint occurs at either the same horizontal or vertical coordinate as one of the points the bisector lies between (so A_1, Z, or an extreme point of $\mathcal{B}$ if an additive bisector part). Let us note, however, that the breakpoints of an additive bisector cannot intersect an edge of $\mathcal{B}$ at the same horizontal or vertical as the active point of $\mathcal{B}$ else this would contradict the convexity of $\mathcal{B}$. Therefore, for any point (b_x, b_y) on an edge of $\mathcal{B}$, the breakpoint of the bisector $B^{\mathcal{B}}(A_1, Z)$ will hit this point only if the distance $l_1^{\mathcal{B}}((b_x, b_y), Z)$ is identical to $l_1^{\mathcal{B}}((b_x, b_y), A_1)$ and

- (b_x, b_y) is visible from Z and exclusively $x = b_x$ or $y = b_y$, or
- (b_x, b_y) is visible from A_1 and exclusively $b_x = 0$ or $b_y = 0$

(note that it only needs to be visible from the facility on whose horizontal or vertical the breakpoint is lying to include additive bisectors).

Therefore, for $b_x = 0$ or $b_y = 0$ (so only for particular barriers), (b_x, b_y) is a site of intersection (without the active extreme point of $\mathcal{B}$ having this coordinate) if and only if

$$l_1^{\mathcal{B}}((b_x, b_y), Z) = |b_y|$$

with $sgn(y) = sgn(b_y)$ for $b_x = 0$, or

$$l_1^{\mathcal{B}}((b_x, b_y), Z) = |b_x|$$

with $sgn(x) = sgn(b_x)$ for $b_y = 0$, which are geodesic diamonds centred at (b_x, b_y).

When considering intersections at $b_x = x$ or $b_y = y$ for a fixed point (b_x, b_y) (so this breakpoint is visible from Z) it must be the case that Z lies on $(b_x, b_y \pm l_1^{\mathcal{B}}((b_x, b_y), A_1))$ or $(b_x \pm l_1^{\mathcal{B}}((b_x, b_y), A_1), b_y)$, depending on the side of $\mathcal{B}$ and the position of A_1 in relation to it. Our new partition lines are composed of these points so, for any adjacent vertices B_i and B_j of $\mathcal{B}$, the lines in question, since we can represent the edge as $y = \frac{b_{j_y} - b_{i_y}}{b_{j_x} - b_{i_x}} x - \frac{b_{i_x} b_{j_y} - b_{i_x} b_{i_y} - b_{j_x} b_{i_y} + b_{i_x} b_{i_y}}{b_{j_x} - b_{i_x}}$, are

$$\begin{aligned}
y &= \frac{b_{j_y} - b_{i_y}}{b_{j_x} - b_{i_x}} x - \frac{b_{i_x} b_{j_y} - b_{i_x} b_{i_y} - b_{j_x} b_{i_y} + b_{i_x} b_{i_y}}{b_{j_x} - b_{i_x}} \\
&\pm l_1^{\mathcal{B}} \left(\left(x, \frac{b_{j_y} - b_{i_y}}{b_{j_x} - b_{i_x}} x - \frac{b_{i_x} b_{j_y} - b_{i_x} b_{i_y} - b_{j_x} b_{i_y} + b_{i_x} b_{i_y}}{b_{j_x} - b_{i_x}} \right), A_1 \right), \\
x &= \frac{b_{j_x} - b_{i_x}}{b_{j_y} - b_{i_y}} y + \frac{b_{i_x} b_{j_y} - b_{i_x} b_{i_y} - b_{j_x} b_{i_y} + b_{i_x} b_{i_y}}{b_{j_y} - b_{i_y}} \\
&\pm l_1^{\mathcal{B}} \left(\left(\frac{b_{j_x} - b_{i_x}}{b_{j_y} - b_{i_y}} y + \frac{b_{i_x} b_{j_y} - b_{i_x} b_{i_y} - b_{j_x} b_{i_y} + b_{i_x} b_{i_y}}{b_{j_y} - b_{i_y}}, y \right), A_1 \right)
\end{aligned}$$

respectively. These lines are necessary since crossing them will cause a breakpoint to be created or removed, altering the Voronoi diagram as shown in Figure 20.

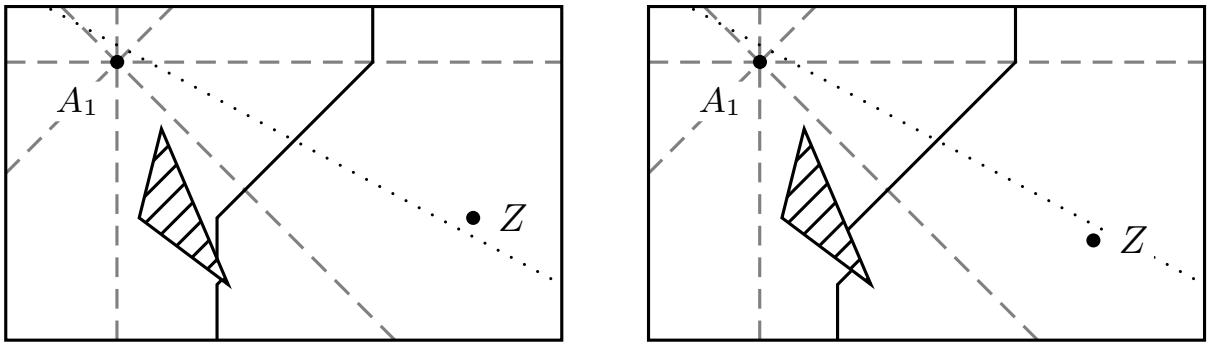


Figure 20: Two placements of $Z = (x, y)$ within $\mathcal{P} = R((-1, -2.5), (4, 0.5))$ with one existing facility $A_1 = (0, 0)$ and a barrier $\mathcal{B} = \Delta((0.2, -1.4), (0.4, -0.6), (1, -2))$ and breakpoint intersection line $7y = -13x + 38$.

The results are analogous for the intersection of breakpoints and the remaining edges of $\mathcal{VD}$ too, but for these we must also now consider the intersection of the breakpoint of an additive bisector part. This does not require much more thought. These breakpoints may only occur on a non-barrier edge of $\mathcal{VD}$ that is at the same x– or y–coordinate of the extreme vertex of $\mathcal{B}$ which is acting as an active point of the additive bisector.

We will phrase these lines in the context of an edge between adjacent vertices P_i and P_j of $\mathcal{P}$, but the same is true for any edge of $\mathcal{VD}$. A breakpoint of the additive bisector will occur, as before, on a

horizontal or vertical line through one of the active points. We have calculated those partition lines for breakpoints sharing a horizontal or vertical with an active point that is a facility. Now we must do so for the extreme vertex which becomes an active point of the additive bisector. For an extreme vertex $B^* = (b_x^*, b_y^*)$, if we have an edge of $\mathcal{P}$ between adjacent vertices P_i and P_j that intersects, visibly, either $x' = b_x^*$ or $y' = b_y^*$, then there may be an intersection between the breakpoint of the bisector and the edge. These intersection points are $\left(b_x^*, \frac{(p_{iy}-p_{jy})b_x^*+p_{ix}p_{jy}-p_{iy}p_{jx}}{p_{ix}-p_{jx}}\right)$ and $\left(\frac{(p_{ix}-p_{jx})b_y^*+p_{iy}p_{jx}-p_{ix}p_{jy}}{p_{iy}-p_{jy}}, b_y^*\right)$ respectively and an intersection will occur only if

$$l_1^{\mathcal{B}}(Z, P^*) = l_1^{\mathcal{B}}(A_1, P^*)$$

where P^* is one of the two points identified above. Here we have a new geodesic diamond centred on P^* of radius $l_1^{\mathcal{B}}(A_1, P^*)$ for every intersection of the quadrant lines of an extreme point of $\mathcal{B}$ and an edge of $\mathcal{P}$ that is visible from the extreme point.

Each existing facility A_k will give as many of the first lines (geodesic diamonds about A_k for non-additive breakpoints on quadrant lines of A_k) as the number of times their quadrant lines intersect (visibly) an edge of $\mathcal{VD}$. Since $\mathcal{B}$ is convex, this amounts to a maximum of two partial geodesic diamonds per facility (only one visible intersection of each horizontal and vertical possible), and $\mathcal{P}$ can do this at most four times (since $A_k \in \mathcal{P}$ and $\mathcal{P}$ is convex). We must consider the bisector edges of $\mathcal{VD}$ a little differently by observing that the bisector between A_k and Z will only exist in the Voronoi diagram within $V(A_k)$. Thus the only bisector edges of $\mathcal{VD}$ that can produce one of these partial geodesic diamonds are those on the perimeter of $V(A_k)$, so this can happen a maximum of four times, analogously to the reasoning for $\mathcal{P}$. We also get a line (for non-additive breakpoints on quadrant lines of Z) for every edge of $\mathcal{VD}$. Next, we turn to the additive bisector breakpoint intersections, i.e. the geodesic diamonds about intersection of edges and quadrant lines of extreme vertices. A new partition diamond is created each time $\mathcal{VD}$ intersects a visible quadrant half-line (i.e. does not travel through $\mathcal{B}$) from an extreme vertex of $\mathcal{B}$. There are a maximum of twelve such quadrant half-lines and $\mathcal{P}$ will intersect each of these while the remaining $\mathcal{O}(n)$ edges of $\mathcal{VD}$ can intersect all of them a maximum of $2(n-1)$ times (this is the case where each Voronoi cell crosses each quadrant half-line, and the boundary between two neighbouring Voronoi cells cannot cross more than once). Therefore, by Lemma 4.5, we produce at most $\mathcal{O}(m+n+p)$ lines.

4.2.6 Kink lines

The next category of lines forms part of the diamonds described above but, rather than defining where the bisectors will intersect a new edge, they cause quite a different change, and so deserve their own section.

Recall the corners C_1, C_2, C_3, and C_4 of $Shadow(\mathcal{B})$. Now the kink lines of C_1, C_2, C_3, and C_4 are defined to be the set of $Z = (x, y)$ such that

$$l_1^{\mathcal{B}}(Z, C_k) = l_1^{\mathcal{B}}(A_i, C_k)$$

(alternatively $\diamond_g^{C_k}(A_i)$ – identical to the first case of additive configuration lines but later we shall see that they are different). For C_k visible from both A_1 and Z these lines are, respectively for C_k,

$$y = x - \min_i b_{ix} - |\min_i b_{ix}| + \max_i b_{iy} - |\max_i b_{iy}|,\ x \in \left(\min_i b_{ix}, \min_i b_{ix} + |\min_i b_{ix}| + |\max_i b_{iy}|\right)$$

$$y = -x + \max_i b_{ix} - |\max_i b_{ix}| + \max_i b_{iy} - |\max_i b_{iy}|,\ x \in \left(\max_i b_{ix} - |\max_i b_{ix}| - |\max_i b_{iy}|, \max_i b_{ix}\right)$$

$$y = x - \max_i b_{ix} + |\max_i b_{ix}| + \min_i b_{iy} + |\min_i b_{iy}|,\ x \in \left(\max_i b_{ix} - |\max_i b_{ix}| - |\min_i b_{iy}|, \max_i b_{ix}\right)$$

$$y = -x + \min_i b_{ix} + |\min_i b_{ix}| + \min_i b_{iy} + |\min_i b_{iy}|,\ x \in \left(\min_i b_{ix}, \min_i b_{ix} + |\min_i b_{ix}| + |\min_i b_{iy}|\right).$$

The effect of crossing these kink lines is shown in Figure 21. This could be described as a breakpoint

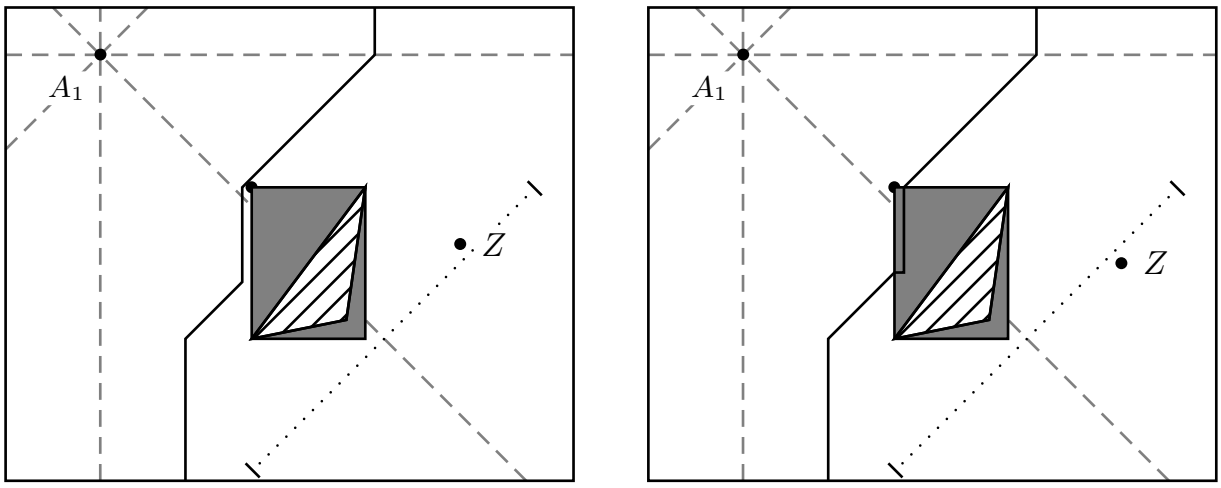


Figure 21: Two placements of Z within $\mathcal{P} = R((-1, -4.5), (5, 0.5))$ with one existing facility $A_1 = (0, 0)$ and $\mathcal{B} = \Delta((2.8, -1.4), (1.6, -3), (2.6, -2.8))$.

hitting $Shadow(\mathcal{B})$ but the interactions are far more complex across these lines than the breakpoint lines. Consider the situation where Z moves so that the bisector touches the corner of $Shadow(\mathcal{B})$ into an area that is not visible to one of the facilities and not already entered. In that case the point of intersection of the clockwise and anticlockwise bisector parts enters $Shadow(\mathcal{B})$ (this must happen since a bisector part cannot leave $Shadow(\mathcal{B})$ once it has entered to intersect elsewhere – it will terminate at $\mathcal{B}$ or the other bisector part). It is because these bisector parts intersect within $Shadow(\mathcal{B})$ that we get a 'kink' since it is only here that the bisector parts can meet as a vertical and a horizontal.

Note that for these kink bisectors to occur we require the interior of the corner to be non-visible from one of the facilities and for the other facility to lie within the quarter plane opposite to $\mathcal{B}$ formed by the extreme lines through the corner. This is due to the fact that we require both extreme points of $\mathcal{B}$ adjacent to the corner to act as an active point in each bisector part for one facility. This forces the other facility to access one active point clockwise about $\mathcal{B}$, and the other anticlockwise, from both active points (not acting through another barrier vertex) and outside the extreme lines in order to get a horizontal and vertical bisector part. Therefore these kink lines occur, given A_i, only for corners C_k for which the quarter plane opposite to or containing $\mathcal{B}$, formed by the extreme lines through C_k, contains A_i (and if A_i is in the quarter plane containing $\mathcal{B}$ then it must not be in the sub-area of $Shadow(\mathcal{B})$ containing C_k). Moreover, we need only the line segments of the geodesic diamonds contained solely within the opposite quarter plane (defined by the extreme lines through C_k) to A_i. The observant reader might wonder how these kink lines are any different to the first case of additive configuration lines. The answer lies exactly in these conditions discussed: while the two facilities in the first case of additive configuration lines must lie within the same quarter plane at C_k, the facilities in the kink construction must inhabit opposite quarter planes.

Since for any A_i there are a maximum of two corners of $Shadow(\mathcal{B})$ at which these kinks can occur, and each geodesic diamond can contribute a maximum of two line segments (by the same argument as the first case of additive configuration lines), we then simply obtain a maximum of four line segments for each existing facility so a total contribution of $\mathcal{O}(n)$ lines.

4.2.7 Wrap-around lines

The final category is an extension of what we can see happening in the previous section. The kink facing A_k formed by crossing the lines described above will 'grow' the further Z travels away from the kink line. Once this touches $\mathcal{B}$ the connected bisector will separate into two parts, and thus the representation of the Voronoi cell will change.

The tip of the kink will touch a specific point B' on an edge of $\mathcal{B}$ if the distances from B' to Z

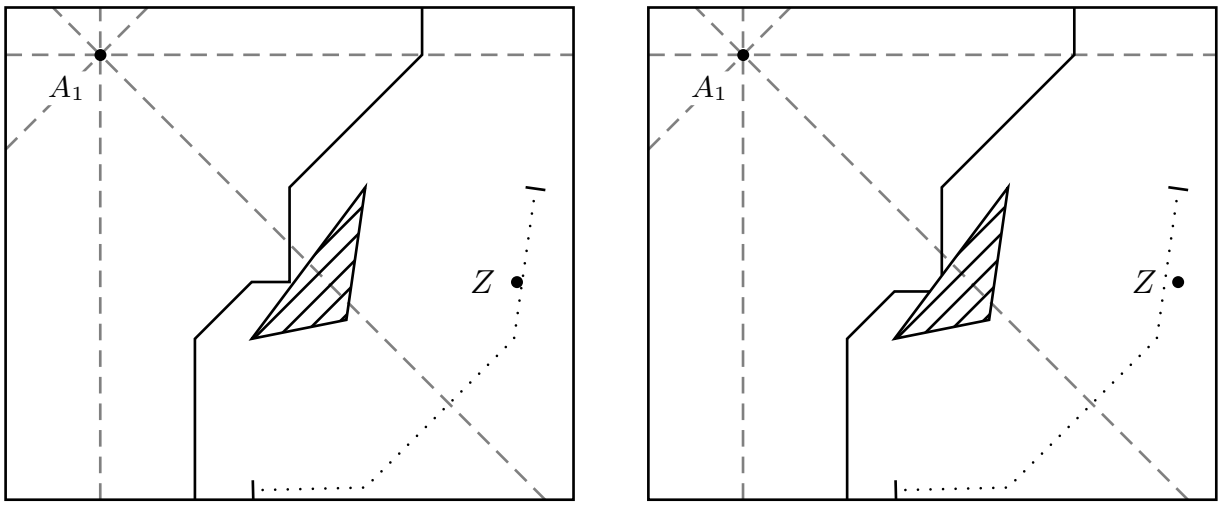


Figure 22: Two placements of Z within $\mathcal{P} = R((-1,-4.7),(5,0.5))$ with one existing facility $A_1 = (0,0)$ and $\mathcal{B} = \Delta((2.8,-1.4),(1.6,-3),(2.6,-2.8))$, showing the wrap-around line for $\overline{(2.8,-1.4)(1.6,-3)}$.

clockwise and anticlockwise about $\mathcal{B}$ are the same and are equal to $l_1^{\mathcal{B}}(B', A_1)$. Thus, for any given (b_x, b_y) on an edge of $\mathcal{B}$, we have $Z = (x, y)$ satisfying

$$l_1^{\mathcal{B},+}(Z,(b_x,b_y)) = l_1^{\mathcal{B},-}(Z,(b_x,b_y)) = l_1(A_1,(b_x,b_y))\,,$$

three equations in two unknowns so solvable (remembering from the previous discussion that, for this kink, A_1 must be visible from the corner of $Shadow(\mathcal{B})$ being crossed and outside the extreme lines of $\mathcal{B}$).

This formulation relies on the placement of Z in relation to $\mathcal{B}$ (and where $\mathcal{B}$ is located in relation to A_1) – recall that the point of intersection must be non-visible from Z. We will show a worked example for Z restricted to $x \geq \max b_{xi}$ and $\min b_{y_i} \leq y \leq \max b_{y_i}$, assuming that the edge between B_i and B_j is in the fourth quadrant of A_1, visible from A_1. In this instance, for (b_x, b_y) lying on $y' = \frac{b_{y_j}-b_{y_i}}{b_{xj}-b_{xi}}x' + \frac{b_{xj}b_{y_i}-b_{xi}b_{y_j}}{b_{xj}-b_{xi}}$ (the edge between B_i and B_j), $l_1^{\mathcal{B},+}(Z,(b_x,b_y)) = x + b_x - 2\min b_{xi} + y + b_y - 2\min b_{y_i}$, $l_1^{\mathcal{B},-}(Z,(b_x,b_y)) = x - b_x - y - b_y + 2\max b_{y_i}$, and $l_1^{\mathcal{B}}(A_1,(b_x,b_y)) = b_x - b_y$. Since the line must satisfy $l_1^{\mathcal{B},+}(Z,(b_x,b_y)) = l_1^{\mathcal{B},-}(Z,(b_x,b_y))$ we have

$$b_x = -y - b_y + \min b_{xi} + \min b_{y_i} + \max b_{y_i}\,.$$

Putting this into the equation of the edge between B_i and B_j we get

$$\begin{aligned} b_y &= \frac{b_{y_j}-b_{y_i}}{b_{xj}-b_{xi}}\left(-y - b_y + \min b_{xi} + \min b_{y_i} + \max b_{y_i}\right) + \frac{b_{xj}b_{y_i}-b_{xi}b_{y_j}}{b_{xj}-b_{xi}} \\ &\Rightarrow (b_{xj}-b_{xi}+b_{y_j}-b_{y_i})b_y = -(b_{y_j}-b_{y_i})y + (b_{y_j}-b_{y_i})(\min b_{xi} + \min b_{y_i} + \max b_{y_i}) + b_{xj}b_{y_i} - b_{xi}b_{y_j}\,. \end{aligned}$$

Now since $l_1^{\mathcal{B},+}(Z,(b_x,b_y)) = l_1^{\mathcal{B}}(A_1,(b_x,b_y))$ we have

$$b_y = \min b_{xi} + \min b_{y_i} - \frac{x+y}{2}\,.$$

Combining these two gives us our line

$$\begin{aligned} &(b_{xj}-b_{xi}+b_{y_j}-b_{y_i})(\min b_{xi} + \min b_{y_i} - \frac{x+y}{2}) \\ &\qquad = -(b_{y_j}-b_{y_i})y + (b_{y_j}-b_{y_i})(\min b_{xi} + \min b_{y_i} + \max b_{y_i}) + b_{xj}b_{y_i} - b_{xi}b_{y_j} \\ &\Rightarrow y = \frac{b_{xj}-b_{xi}+b_{y_j}-b_{y_i}}{b_{y_j}-b_{y_i}-b_{xj}+b_{xi}}x + 2\frac{(b_{y_j}-b_{y_i})\max b_{y_i} - (b_{xj}-b_{xi})(\min b_{xi} + \min b_{y_i}) + b_{xj}b_{y_i} - b_{xi}b_{y_j}}{b_{y_j}-b_{y_i}-b_{xj}+b_{xi}}\,. \end{aligned}$$

This line contributes to the rightmost line segment in Figure 22 and the other segments are found similarly (using the appropriate representation of the shortest distance). These wrap-around lines give us a maximum of three line segments per edge since the line will only alter when an extreme line of $\mathcal{B}$ is crossed (thereby affecting the calculation of the shortest distance around $\mathcal{B}$). These lines take a variety of gradients as shown in Figure 23.

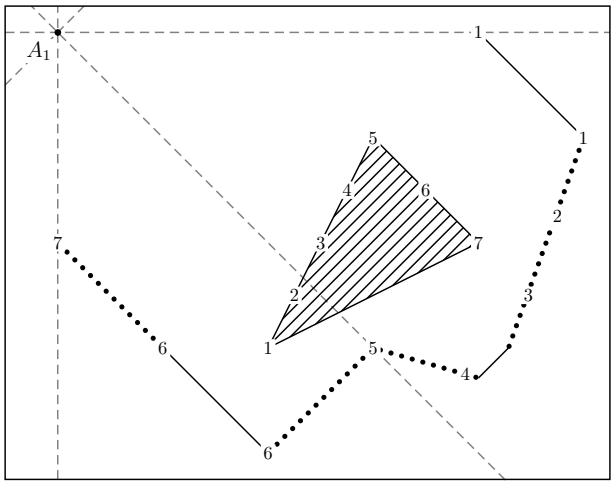


Figure 23: Possible placements of Z within $\mathcal{P} = R((-1,-6),(7,1))$ with one existing facility $A_1 = (0,0)$ and $\mathcal{B} = \Delta((20,-15),(10,-20),(15,-10))$ satisfying $l_1^{\mathcal{B},+}(Z,(b_x,b_y)) = l_1^{\mathcal{B},-}(Z,(b_x,b_y)) = l_1(A_1,(b_x,b_y))$ where the points (b_x,b_y) are specified on $\mathcal{B}$.

A particularly remarkable interaction occurs when a point on $\mathcal{B}$ is equidistant clockwise and anticlockwise around $\mathcal{B}$ to a corner of $Shadow(\mathcal{B})$. In Figure 23 this occurs with point 1, somewhere between points 3 and 4, and point 6. There is not a unique position of Z at which the kink touches these points (as shown by the solid diagonal line). This is further exemplified when every point of the edge is the same distance away from A_1 (for example a diagonal facing A_1). As shown in Figure 24, the position of Z at which the kink intersects a specific part of $\mathcal{B}$ is not unique. It is in fact a diagonal line contained in $Shadow(\mathcal{B})$. This means that every point within the areas of $Shadow(\mathcal{B})$ not containing the diagonal facing A_1 forms a kink on this diagonal. While this does complicate the issue as we now have a kink block rather than a kink line (so the partitioning lines are the appropriate perimeter lines of $Shadow(\mathcal{B})$), it does mean that there may not be a unique optimal solution since multiple locations for Z give the same bisector (with A_1 at least).

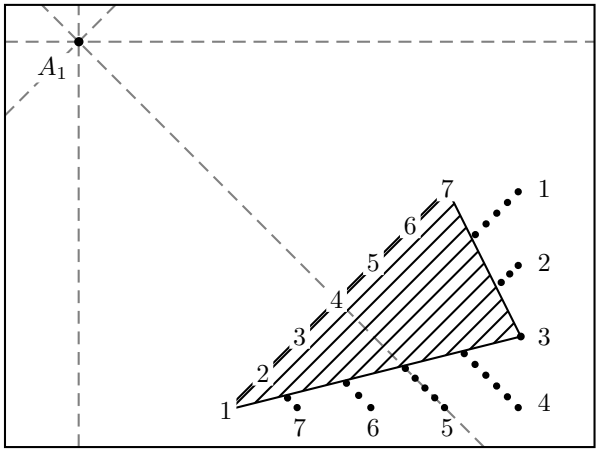


Figure 24: Possible placements of Z within $\mathcal{P} = R(-1,-5.5),(7,0.5))$ with one existing facility $A_1 = (0,0)$ and $\mathcal{B} = \Delta((5,-2),(6,-4),(2,-5))$ where the kink intersects the specified points on $\mathcal{B}$.

Finally there are also wrap-around lines where the shortest paths each way around $\mathcal{B}$ from A_1 are equal. Since A_1 is fixed these are much simpler (being just one line segment), and they can only occur when a corner $C_i = (C_x, C_y)$ of $Shadow(\mathcal{B})$ is equidistant from A_1 both ways around $\mathcal{B}$. The kink will also only form along the diagonal line through C_i so we need no longer ask where the kink will intersect, but merely when. An example of this is shown in Figure 25.

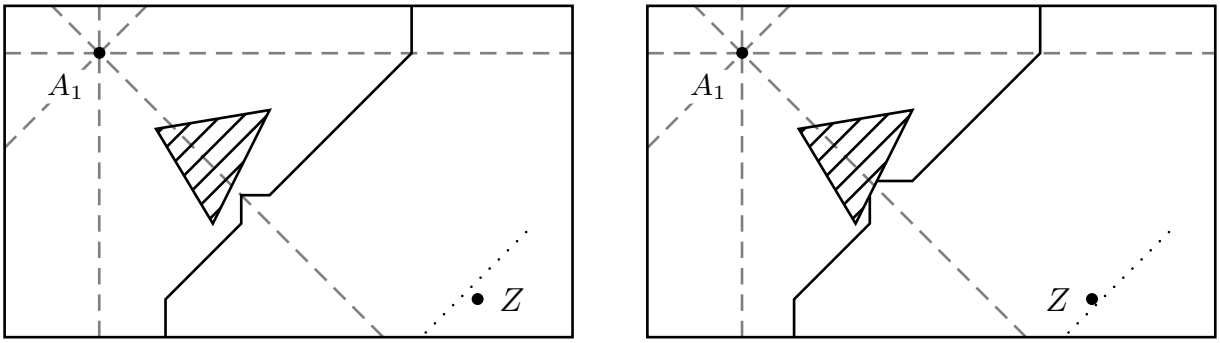


Figure 25: Two placements of Z within $\mathcal{P} = R((-1,-3),(5,0.5))$ with one existing facility $A_1 = (0,0)$ and $\mathcal{B} = \Delta((1.8,-0.6),(1.2,-1.8),(0.6,-0.8))$.

In this case, for the unique point (b_x, b_y) on $\mathcal{B}$ that is the first point of intersection of $\mathcal{B}$ on the same diagonal as C_i, the bisector intersects (b_x, b_y) if and only if Z lies on the line connecting $(C_x, C_y + sgn(C_y - b_y)(l_1^{\mathcal{B}}(A_1,(b_x,b_y)) - l_1(C_i,(b_x,b_y)))$ and $(C_x + sgn(C_x - b_x)(l_1^{\mathcal{B}}(A_1,(b_x,b_y)) - l_1(C_i,(b_x,b_y)), C_y)$. That is

$$y = C_y + sgn(C_y - b_y)(l_1^{\mathcal{B}}(A_1,(b_x,b_y)) - l_1(C_i,(b_x,b_y)) - \frac{sgn(C_y - b_y)}{sgn(C_x - b_x)}(x - C_x)\,.$$

A barrier will produce a maximum of $p+n+8$ line segments as follows. Each edge can only interact with the closest facility A_i to obtain a wrap-around line so the Voronoi diagram $\mathcal{VD}$ partitions the perimeter of $\mathcal{B}$ into at most $p+n$ edge segments. In turn each segment gives a wrap-around line with the (eight) intersections of these lines with the extreme lines of $\mathcal{B}$ causing another line segment to be created, as we have seen.

These being the last partitioning lines, we have a final total of a maximum number of lines to the order of $\mathcal{O}(m+n+p)$ in our partition.

4.2.8 The complete partition and degeneracy

We have introduced all of the partition lines which, from their algebraic formulation, are clearly unique, and have justified why they are necessary. What is still required is to prove that these seven sets of lines are sufficient to partition the space into cells within which the Voronoi diagram is structurally identical, and this is done in Appendix A.4. This leads us to our major result.

Theorem 4.1. *The set of configuration lines, additive configuration lines, extreme lines, geodesic diamonds, breakpoint intersection lines, kink lines, and wrap-around lines described above induces a partition of $\mathcal{P}$ into $\mathcal{O}(m+n+p)$ structurally identical cells.*

Before moving on we have to address the possibility that the bisector produces a quarter plane and how we overcome the issues that this may bring. Quarter planes can exist in bisectors in l_1, even without the inclusion of a barrier's interaction, if the two facilities are on the same diagonal. They may also occur in additive bisectors as shown in Figure 12. While we may comment that in practice this case need not occur since one can appeal to measurement error, for completeness we must still choose how to deal with such a bisector. Does the original facility maintain the custom, or do the customers fall for the attraction of what's shiny and new?

The choice is made simpler by the realisation that we require an exact solution for our algorithm. If it is decided that the original facility keeps the demand in the quarter plane then one can easily see that we could have the case where, within a cell of structural identity, the area of $V(Z)$ will increase as it approaches the point on one of the edges at which a quarter plane would be created. This maximisation would therefore have a supremum, but not a maximum as we require. Therefore we do require the demand within the quarter plane to go to the new facility. One can interpret this as meaning that our customers are curious creatures, intrigued by novelty.

5 Various Conditional Facility Location Problems

The partition derived in the previous section can be utilised to solve the five conditional planar facility location problems formulated earlier. We first list them below, recapitulating their objectives in the Voronoi diagram environment, before investigating exact polynomial algorithms for solving each one.

The Conditional Median and Antimedian Problem (CMP and CAMP):

$$F_{CM}(Z) = \sum_{i=1}^{n+1} \int \int_{Q=(u,v)\in V(A_i)} l_1^{\mathcal{B}}(A_i, Q) du dv$$

The Conditional Centre Problem (CCP):

$$F_{CC}(Z) = \max_{i=1,\ldots,n+1} \max_{Q=(u,v)\in V(A_i)} l_1^{\mathcal{B}}(A_i, Q)$$

The Market Share Problem (MSP):

$$F_{MS}(Z) = \sum_{i=k+1}^{n+1} \int \int_{Q=(u,v)\in V(A_i)} 1 du dv$$

The Conditional Maximal Covering Location Problem (CMCLP):

$$F_{MCL}(Z) = \sum_{i=1}^{n+1} \int \int_{Q=(u,v)\in V(A_i)\cap B_R(A_i)} 1 du dv$$

5.1 Solving the Conditional Median and Antimedian Problems

The Conditional Median Problem requires the calculation of a more difficult integral. This is evaluated in Averbakh et al. (2015) using centroid triangulation, which is shown in Figure 26, and the notion of quadrant identity. However, in Averbakh et al. (2015) the Voronoi cells were always star-shaped while the inclusion of a barrier easily violates this property, as seen in Figure 27a (adapted from Figure 14b). Another issue is how the shortest path must be considered around $\mathcal{B}$, and so even some star-shaped Voronoi cells (taking out $\mathcal{B}$) need careful consideration, as shown in Figure 27b (taken from Figure 21a).

Note that the shortest paths from each facility to each vertex of its Voronoi cell have been included in Figure 27, and for this in some cases we must include the lines of equidistance, shown in Figure 27[right] as the dotted line (this line is always piecewise linear). We include this in our subdivision as an edge of $V(A_k)$, thereby separating the areas of points reached travelling clockwise and anticlockwise about $\mathcal{B}$.

Whilst these Voronoi cells cannot be broken down using centroid triangulation, an adaptation of this method is easily found. To solve the issue we simply subdivide the cell into easily manageable pieces using the quadrant lines through the vertices of $\mathcal{B}$. For centroid triangulation to work we need the area to be visible in l_2 from the facility. However, a vertex of $\mathcal{B}$ could easily block the line of sight (this need not always be an extreme vertex as we have explored before, as Euclidean visibility is reduced by a barrier to a much greater extent than l_1 visibility). Therefore we divide the Voronoi cell into subcells, each one designated to a vertex of $V(A_k)$ which could block l_2-sight from the facility and

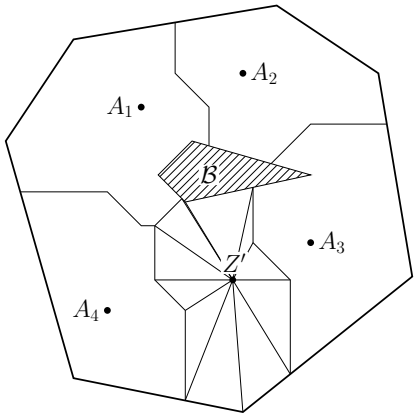


Figure 26: Centroid triangulation of a Voronoi cell $V(Z)$.

so become the acting centroid for the l_2-non-visible area of $V(A_k)$ – an additive centroid triangulation, if you will.

Consider a vertex of $V(A_k)$ on a reflex angle l_2-visible from A_k (these are the only vertices which could possibly interrupt line of sight). For this vertex we partition $V(A_k)$ using the quadrant half-line from the vertex which is l_2-visible from A_k and travelling away from A_k (only one of which satisfies these conditions, unless the vertex and A_k are on the same horizontal or vertical lines in which case we ignore this vertex). Once these lines have been found, the partition cell of $V(A_k)$ containing A_k will be l_2-visible from A_k and so centroid triangulation will work and give the desired integral.

The remaining partition cells of $V(A_k)$ can be treated the same, with the blocking vertex acting in the place of A_k. We partition the cells further into smaller cells in which we can use centroid triangulation until we are left with a cell which is wholly l_2-visible from the active vertex of the cell. Examples of this method are shown in Figure 28. Importantly, the shortest path between every point in each partition cell and A_k can pass by the acting centroid vertex of the cell (and subcells) in which it is contained. So, for every point Q in centroid triangle T in the partition cell of blocking vertex B', we have $l_1^{\mathcal{B}}(A_i, Q) = l_1^{\mathcal{B}}(A_i, B') + l_1^{\mathcal{B}}(B', Q)$. Thus the integral we require in this partition cell is

$$\int_{Q\in T} l_1^{\mathcal{B}}(A_i, Q)dQ = \int_{Q\in T} l_1^{\mathcal{B}}(B', Q)dQ + \int_{Q\in T} l_1^{\mathcal{B}}(A_i, B')dQ = \int_{Q\in T} l_1(B', Q)dQ + l_1^{\mathcal{B}}(A_i, B')Area(T)\,.$$

Therefore we can calculate the objective function for the Conditional Median Problem and find its maximiser, just as in the barrier-free case.

The representation of this partition of $V(A_k)$ will only change when Z crosses a quadrant line through one of these reflex vertices (thereby changing which quadrant half-line from this vertex is

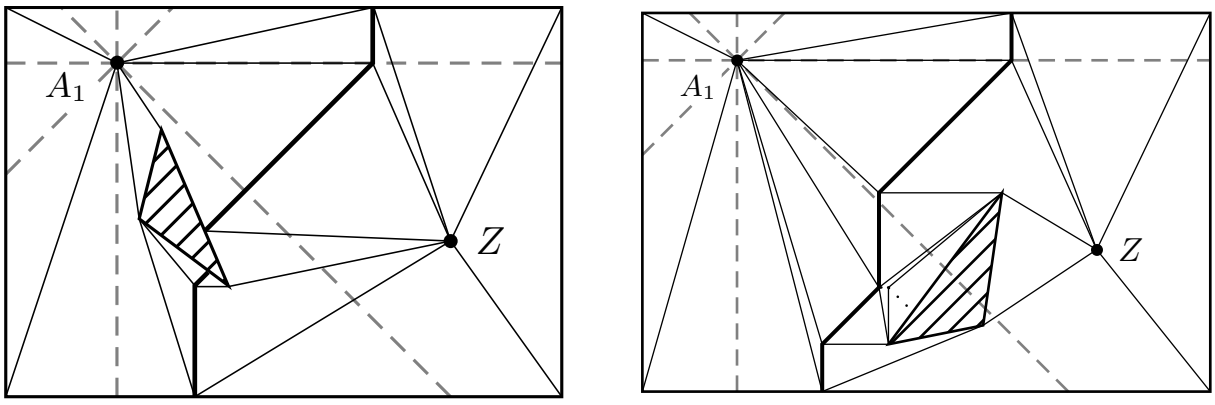


Figure 27: Two examples of non-star-shaped Voronoi cells.

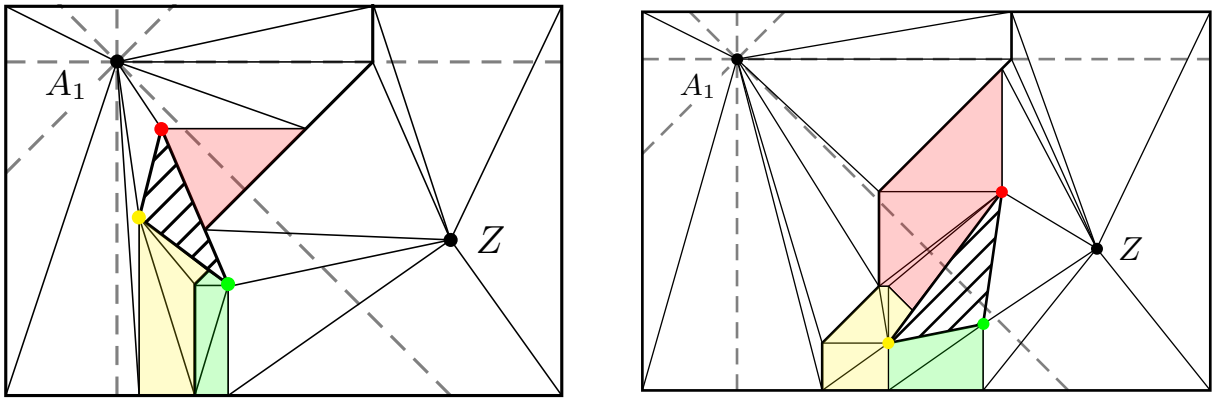


Figure 28: Two examples of additive centroid triangulation with blocking vertices and their partition cells colour-coded.

chosen to divide $V(A_k)$). These reflex vertices can only be vertices of $\mathcal{B}$ so this produces $2p$ lines. We must also have at most $2m$ lines from the quadrant lines of the vertices of $\mathcal{P}$ required for quadrant identity (the notion of which we have extended through our additive centroid triangulation), which is needed to ensure that the structure of the triangulation does not change. This leaves us with $\mathcal{O}(m+n+p)$ partition lines of $\mathcal{P}$, creating at most $\mathcal{O}((m+n+p)^2)$ partition cells.

Regarding the algorithmic complexities, the barrier-restricted Voronoi diagram for the existing facilities can be calculated in $\mathcal{O}((n+p)\log^2(n+p))$ time (Mitchell, 1992) and its overlay with the market region $\mathcal{P}$ can be calculated in $\mathcal{O}(m+n+p)$ time (Finke & Hinrichs, 1995). Finally, as the objective is again cubic in x and y within each partition cell, from Averbakh et al. (2015), the minima in the interior of the cell and the minimum on any edge of the cell take a constant time to be found, so we obtain the following result.

Theorem 5.1. *The Conditional Median Problem can be solved in $\mathcal{O}((m+n+p)^2)$ time.*

The Conditional Antimedian Problem is addressed in Averbakh et al. (2015) by simply including all bounding lines of the l_1-disc of the specified radius about each A_i. We complete the same procedure in the barrier-constrained problem, except this time the l_1-disc is the geodesic diamond. This contributes at most $7n$ lines so the number of cells in the partition is again $\mathcal{O}((m+n+p)^2)$, giving us the following:

Theorem 5.2. *The Conditional Antimedian Problem can be solved in $\mathcal{O}((m+n+p)^2)$ time.*

5.2 Solving the Conditional Centre Problem

The solution to the Conditional Centre Problem can now also be found using the same reasoning as that offered in Averbakh et al. (2015), which relies solely on the vertices for the Voronoi cells being linear in the coordinates of Z. Using the fact that the number of vertices in the Voronoi diagram is $\mathcal{O}(m+n+p)$, an identical argument to that in Averbakh et al. (2015) gives:

Theorem 5.3. *The Conditional Centre Problem can be solved in $\mathcal{O}((m+n+p)^3)$ time.*

5.3 Solving the Market Share Problem

The objective function for the Market Share Problem is exactly the area of the Voronoi cells of the non-competitor facilities. Therefore we require the ability to calculate the area of a Voronoi cell. Remembering that $\mathcal{B}$ removes demand, we observe that $\mathcal{B}$ can be contained either entirely within $V(Z)$ or not at all (since $V(Z) \cap int(\mathcal{B}) = \emptyset$). First suppose the latter. Since $V(Z)$ is simple and polygonal, if the boundary of $V(Z)$ has vertices $D_1, \ldots, D_j$ then, from Gauss's area formula,

$$Area(V(Z)) = \sum_{i=1}^{j} \left(D_i^x \, D_{i+1}^y \; - \; D_{i+1}^x \, D_i^y \right).$$

The representations of the coordinates and of the line segments of the bisectors are clearly linear in the coordinates of the facilities to which the bisector belongs, so the vertices $D_1, \ldots, D_j$ are linear in the coordinates of Z. Thus $Area(V(Z))$ is a quadratic function in the coordinates of Z and can easily be maximised over its partition cell.

Alternatively, if $\mathcal{B}$ (with vertices $B_1, \ldots, B_l$) is contained entirely within $V(Z)$ then

$$Area(V(Z)) = \sum_{i=1}^{j} \left(D_i^x \, D_{i+1}^y \; - \; D_{i+1}^x \, D_i^y \right) - \sum_{i=1}^{l} \left(B_i^x \, B_{i+1}^y \; - \; B_{i+1}^x \, B_i^y \right)$$

which is likewise clearly quadratic and easily maximised.

Thus, using similar arguments to those in the Conditional Median Problem we obtain:

Theorem 5.4. *The Market Share Problem can be solved in $\mathcal{O}((m+n+p)^2)$ time.*

5.4 Solving the Conditional Maximal Covering Location Problem

The objective in the Conditional Maximal Covering Location Problem can be calculated by the same logic since $V(A_i) \cap B_R(A_i)$ remains simple and polygonal. However, while structural identity maintains the algebraic representation of $V(A_i)$, it is no longer sufficient to obtain an identical representation of $V(A_i) \cap B_R(A_i)$. To solve this, Averbakh et al. (2015) introduced the concept of sphere identity: two objects are sphere identical if they are either structurally identical, identical vertices of $B_R(A_i)$, or similar intersection points of the same edge of $B_R(A_i)$ and the similar segment of $\mathcal{VD}$. Following this they described spherical lines (each one of which ensures that a specific edge of $B_R(Z)$ intersects a specific segment of $B(A_i, Z)$) within the partition of which $V(A_i) \cap B_R(A_i)$ has an identical representation. We can follow their logic, extending the spherical lines to also include the intersection of the coverage dish with $\mathcal{B}$ through our geodesic formulations, introducing $\mathcal{O}(p)$ lines to give $\mathcal{O}(m+n+p)$ spherical lines overall. This leads us to our final theorem.

Theorem 5.5. *The Conditional Maximal Covering Location Problem can be solved in $\mathcal{O}((m+n+p)^2)$ time.*

6 Conclusions

This paper identifies the partition of a convex polygonal market region, within which n facilities are fixed and a polygonal barrier is situated, into convex polygonal cells with the result that the representation of the objective function is preserved for any possible new facility locations in each cell. The partition obtained by Averbakh et al. (2015) was extended by identifying a set of six novel and very dissimilar lines which ensure the preservation of the representation of the objective function by considering the position of the barrier in relation to the bisector from the barrier-free problem and how its interaction with the bisectors affects the representation of the objective function. These lines were found by exploiting a parametric representation of the objective function when considering its influence from the Voronoi diagram. By this means we were able to prove that polynomial exact algorithms exist for solving the facility location problems with barriers.

The derived results are fairly general and can easily be applied to a range of location problems in the rectilinear plane to find exact solutions. The techniques of this paper are efficacious in proving similar results. The methodology demonstrated for the conditional median and the market share problems should be extendable to the majority of problems in which the objective function is additive, i.e. is the integral of all customers' individual contributions, depending solely on each customer's nearest facility and its location in relation to them. The subproblems over the cells of the derived partitions were solved analytically for the conditional median and the market share problems; however, numerical methods may be required for more complex problems.

There are two further classes of models worth mentioning from the literature where the approach proposed in this paper may be beneficial. Conditional location problems occur naturally in time-dynamic models (Okabe et al., 2000). Unlike models where all facilities can be constructed simultaneously, in time-dynamic models facilities are constructed not instantaneously but rather in stages,

perhaps due to considerations such as limited resources. A typical example of a member of this family of problems is the locating of one new facility per time period as presented in Okabe et al. (2000). A standard goal of these problems is to minimise the total sum of travel times between customers and their closest facility over all time periods.

Finally, as in the case of many of the results proven above, the barrier need not always be convex, and it could be investigated whether it need ever be convex. This work on barrier-constrained facility location problems can also naturally be extended to the multi-barrier problem which would then facilitate the relaxation of a convex market region. The relaxation of the uniform demand assumption was already explored in Byrne (2016) in the barrier-free scenario, and so a natural extension would be to combine the two results and investigate the problems in the presence of multiple, non-convex barriers where demand is distributed continuously and non-uniformly over a non-convex market region.

References

Aneja, Y., & Parlar, M. (1994). Algorithms for Weber facility location in the presence of forbidden regions and/or barriers to travel. *Transportation Science*, *28*(1), 70–76.

Averbakh, I., Berman, O., Kalcsics, J., & Krass, D. (2015). Structural properties of Voronoi diagrams in facility location problems with continuous demand. *Operations Research*, *62*(2), 394–411.

Berman, O., & Huang, R. (2008). The minimum weighted covering location problem with distance constraints. *Computers & Operations Research*, *35*(2), 356–372.

Bischoff, M., & Klamroth, K. (2007). An efficient solution method for Weber problems with barriers based on genetic algorithms. *European Journal of Operational Research*, *177*(1), 22–41.

Byrne, T. (2016). *Conditional facility location problems with piecewise polynomial demand* (MSc dissertation). School of Mathematics, The University of Edinburgh.

Canbolat, M. S., & Wesolowsky, G. O. (2012). On the use of the Varignon frame for single facility Weber problems in the presence of convex barriers. *European Journal of Operational Research*, *217*(2), 241–247.

Dearing, P. M., Hamacher, H., & Klamroth, K. (2002). Dominating sets for rectilinear center location problems with polyhedral barriers. *Naval Research Logistics*, *49*, 647–665.

Dearing, P. M., Klamroth, K., & Segars, R. J. (2005). Planar location problems with block distance and barriers. *Annals of Operations Research*, *136*(1), 117–143.

Drezner, Z. (Ed.). (1995a). *Facility location: A survey of applications and methods.* Springer, New York.

Drezner, Z. (1995b). Replacing discrete demand with continuous demand. In Z. Drezner (Ed.), *Facility location: A survey of applications and methods* (chap. 2). Springer, New York.

Drezner, Z., & Hamacher, H. (Eds.). (2002). *Facility location: Applications and theory.* Springer-Verlag, Berlin Heidelberg New York.

Eiselt, H. A., Pederzoli, G., & Sandblom, C. L. (1985). On the location of a new service facility in an urban area. *Proceedings of the Administrative Sciences Association of Canada*, *6*, 356–372.

Erlenkotter, D. (1989). The general market area model. *Annals of Operations Research*, *18*, 45–70.

Fekete, S. P., Mitchell, J. S. B., & Beurer, K. (2005). On the continuous Fermat-Weber problem. *Operations Research*, *53*(1), 61–76.

Finke, U., & Hinrichs, K. H. (1995). Overlaying simply connected planar subdivisions in linear time. In *Proceedings of the eleventh annual symposium on computational geometry* (p. 119-126).

Francis, R., & Lowe, T. (2019). Aggregation in location. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location Science* (pp. 537–556). Springer, Cham.

Francis, R., & White, J. (1974). *Facility layout and locations: An analytic approach.* Prentice Hall, Englewood Cliffs, N.J.

Hakimi, S. (1964). Optimal location of switching centers and the absolute centers and medians of a graph. *Operations Research*, *12*, 450–459.

Hamacher, H. W., & Nickel, S. (1995). Restricted planar location problems and applications. *Naval Research Logistics*, *42*(6), 967–992.

Katz, I., & Cooper, L. (1981). Facility location in the presence of forbidden regions, I: Formulation and the case of Euclidean distance with one forbidden circle. *European Journal of Operational Research*, *6*(2), 166–173.

Klamroth, K. (2001). A reduction result for location problems with polyhedral barriers. *European Journal of Operational Research*, *130*(3), 486–497.

Laporte, G., Nickel, S., & Saldanha da Gama, F. (Eds.). (2019). *Location science.* Springer International Publishing.

Larson, R., & Sadiq, G. (1983). Facility locations with the Manhattan metric in the presence of barriers to travel. *Operations Research*, *31*(4), 652–669.

Launhardt, C.-F. (1900). *The principles of location: The theory of the trace. Part I: The commercial trace.* Lawrence Asylum Press, Madras. (A. Bewley, trans., 1900)

Lösch, A. (1954). *The Economics of Location.* Yale University Press, New Haven. (W.H. Woglom, trans.)

Manne, A. (1964). Plant location under economies of scale-decentralization and computation. *Management Science*, *11*, 213–235.

Marucheck, A., & Aly, A. (1981). An efficient algorithm for the location-allocation problem with rectangular regions. *Naval Research Logistics Quarterly*, *28*, 309–323.

Matisziw, T. C., & Murray, A. T. (2009a). Area coverage maximization in service facility siting. *Journal of Geographical Systems*, *11*(2), 175–189.

Matisziw, T. C., & Murray, A. T. (2009b). Siting a facility in continuous space to maximize coverage of a region. *Socio-Economic Planning Sciences*, *43*(2), 131–139.

Mitchell, J. (1992). L_1 shortest paths among polygonal obstacles in the plane. *Algorithmica*, *8*(1), 55–88.

Murat, A., Verter, V., & Laporte, G. (2011). A multi-dimensional shooting algorithm for the two-facility location-allocation problem with dense demand. *Computers & Operations Research*, *38*, 450–463.

Murray, A. T., O'Kelly, M. E., & Church, R. L. (2008). Regional service coverage modeling. *Computers & Operations Research*, *35*(2), 339–355.

Murray, A. T., & Tong, D. (2007). Coverage optimization in continuous space facility siting. *International Journal of Geographical Information Science*, *21*(7), 757–776.

Newell, F. (1973). Scheduling, location, transportation, and continuum mechanics: some simple approximations to optimization problems. *SIAM Journal of Applied Mathematics*, *25*, 346–360.

Okabe, A., & Aoyagi, M. (1993). Spatial competition of firms in a two-dimensional bounded market. *Regional Science and Urban Economics*, *23*, 259–289.

Okabe, A., Boots, B., Sugihara, K., & Chiu, S. (2000). *Spatial tessellations: Concepts and applications of Voronoi diagrams*. John Wiley and Sons, Chichester, UK.

Okabe, A., & Suzuki, A. (1987). Stability of spatial competition for a large number of firms on a bounded two-dimensional space. *Environment and Planning A*, *19*(8), 1067–1082.

Oğuz, M., Bektaş, T., & Bennell, J. A. (2018). Multicommodity flows and Benders decomposition for restricted continuous location problems. *European Journal of Operational Research*, *266*(3), 851–863.

Suzuki, A., & Drezner, Z. (1996). The p-center location problem in an area. *Location Science*, *4*(1), 69–82.

Weber, A. (1909). *Theory of the Location of Industries*. University of Chicago Press. (C.J. Friedrich, trans., 1929)

Wei, H., Murray, A., & Xiao, N. (2006). Solving the continuous space p-centre problem: planning application issues. *IMA Journal of Management Mathematics*, *17*(4), 413–425.

Weiszfeld, E. (1937). Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Mathematical Journal*, *43*, 335–386.

Wesolowsky, G. (1993). The Weber problem: History and perspective. *Location Science*, *1*, 5–23.

Xue, G., Rosen, J., & Pardalos, P. (1996). A polynomial time dual algorithm for the Euclidean multifacility location problem. *Operations Research Letters*, *18*(4), 201–204.

A Omitted proofs

A.1 Proof of Lemma 4.3

Lemma 4.3. *A bisector's direction will only change by* $45°$ *unless it crosses the line of equidistance around* $\mathcal{B}$ *in which case it can also change by* $90°$.

Proof. It is interesting to note that, when the bisector part between two active points P and Q enters an area of $\mathbb{R}^2$ that is not visible from P but visible from facility Q, it will change its direction by $45°$ (by the continuity of the additive bisector), choosing the bearing that travels equidistant from the vertex of $\mathcal{B}$ most recently passed on the shortest path from P. That is, a horizontal/vertical or a diagonal part will become a diagonal or horizontal/vertical part respectively. Otherwise a bisector part's direction remains unchanged. The only possibility for a different situation to arise would be if the bisector moves from an area visible to both facilities to one non-visible from both, but this is impossible (unless in the degenerate quarter plane case which is addressed later) for a convex $\mathcal{B}$ and convex $\mathcal{P}$ (for convex $\mathcal{P}$ only vertices of $\mathcal{B}$ can become active points). This can be seen by realising that a vertex becomes the new active point when the bisector crosses either the horizontal or vertical through it. If the bisector were to move into an area non-visible from both current active points, two new active points P and Q would have to be found (and at the same vertical or horizontal to the bisector), and of course they must be either side of the bisector. If these active points lie on the same horizontal or vertical as one another, then by the convexity of $\mathcal{B}$ the line PQ must be contained in $\mathcal{B}$, so there is no such area for the bisector to enter. For the other case, suppose that the bisector hits the horizontal of P, and vertical of Q. For P to act as an active point the bisector must travel vertically before hitting the horizontal of P. However, identically for Q the bisector must travel horizontally before hitting the vertical of Q, which is clearly a contradiction.

Therefore a bisector will only change direction by $45°$ unless it is crossing the line of equidistance around $\mathcal{B}$ (where the bisector part being used changes) in which case it can also change by $90°$. We know that the angle cannot ever be $135°$ by noticing that, due to being simply an additive bisector of sorts, the line of equidistance around $\mathcal{B}$ is composed of only horizontal, vertical, and diagonal line segments. Now, suppose for contradiction that there is a change in direction of the bisector of $135°$. As explained above, a bisector part can only change direction by $45°$ so this must be the point at which the upper and lower bisector parts meet (without loss of generality the shortest paths from Q must travel either side of $\mathcal{B}$, say). Therefore the line of equidistance from Q around $\mathcal{B}$ intersects the bisector (travelling from one side of the bisector to the other) at the point of this $135°$ direction change. However, as the line of equidistance enters the Voronoi cell of P containing the acute angle, it must, for however long, travel along the bisector itself since these bisector line segments are themselves a diagonal, horizontal, or vertical. This is impossible. One can see this as, for small enough ε, travelling ε perpendicularly from the line of equidistance will get ε closer to Q, much like moving ε perpendicular to a bisector into a cell will get ε closer to the cell's facility. On the line segment shared by the line of equidistance and the bisector, venturing ε into $V(P)$ perpendicular to these lines will move both ε closer to P and ε closer to Q, and so should also be contained in the bisector. □

A.2 Proof of Lemma 4.4

Lemma 4.4. *For facilities* $A_1, \ldots, A_n$ *and a* p*-sided convex barrier* $\mathcal{B}$ *located with an* m*-sided convex polygon* $\mathcal{P}$*, the number of edges and the number of vertices in* $\mathcal{VD}(A_1, \ldots, A_n)$ *are each* $\mathcal{O}(m+n+p)$.

Proof. For this proof we utilise the dual graph (or the Delaunay triangulation) of the Voronoi diagram. We obtain this by taking the facilities A_1 to A_n as our vertices and we draw an edge between two vertices if their facilities' Voronoi cells neighbour one another. Since each Voronoi cell is connected, this graph is a planar graph (no two edges must cross as this would prevent one of the neighbouring conditions – the edge between A_i and A_j can be drawn through only $V(A_i)$ and $V(A_j)$, allowing curves) with n vertices, and so Euler's formula gives the maximum number of edges of the dual graph as $3n-6$ for $n > 2$, and 1 if $n = 2$.

Each edge of this graph corresponds to a bisector between two facilities. As seen above, the bisector in the barrier-free problem, composed of three line segments, is changed by the addition of the barrier

whenever an extreme vertex of $\mathcal{B}$ becomes an active extreme vertex, and each change causes a shift in the original bisector, or a horizontal or vertical to become a diagonal, or a diagonal to become a horizontal or vertical, increasing the edge count by at most one. The barrier itself can also intersect the bisector once, splitting an edge into two, so creating another edge. There are a maximum of four extreme vertices of $\mathcal{B}$ so the maximum number of edges of a bisector is $3 + 4 + 1 = 8$. Therefore every edge in the dual graph corresponds to a bisector within the Voronoi diagram of at most eight edges.

Finally, the edges of $\mathcal{B}$ and $\mathcal{P}$ must be included in the Voronoi diagram. According to Finke and Hinrichs (1995) the overlay of two simply connected planar subdivisions, with $\mathcal{O}(a)$ and $\mathcal{O}(b)$ edges respectively, has $\mathcal{O}(a + b)$ edges. Therefore there are $\mathcal{O}(m + n + p)$ edges in $\mathcal{VD}(A_1, \ldots, A_n)$. It follows from the fact that the number of vertices cannot exceed the number of edges that we also have $\mathcal{O}(m + n + p)$ vertices in $\mathcal{VD}(A_1, \ldots, A_n)$. □

A.3 Proof of Lemma 4.5

Lemma 4.5. *For a convex $\mathcal{B}$, geodesic diamonds have a maximum of seven edges.*

Proof. Let us first look at geodesic diamonds centred on vertices of $\mathcal{B}$. We will show the result is true by classifying the three cases of vertices of $\mathcal{B}$. A vertex of $\mathcal{B}$ can either be a corner of $Shadow(\mathcal{B})$, lie on one edge of $Shadow(\mathcal{B})$, or be neither (an 'interior' vertex). The possible geodesic diamonds for each of these vertices is shown in Figure 29 where the number of line segments of each diamond is labelled and can be seen never to exceed seven. This is because once the diamond enters a non-visible area of $\mathcal{P}$ it changes direction by a set angle, much like the bisector parts. This corresponds to a new vertex of $\mathcal{B}$ acting as an active point within the diamond (similarly to within additive bisectors). Once the diamond crosses a quadrant line of an active vertex it will then turn by 90°. Therefore as the geodesic diamond expands from the vertex, each time the intersection with $\mathcal{B}$ passes an extreme point it will introduce another line segment, with two produced from a corner vertex (a 'double' extreme point). Finally we observe, as expected, that the smallest number of line segments of a geodesic diamond centred on an interior vertex is three, on an edge vertex is four, and on a corner vertex is five, and the number of extreme points for the geodesic diamond still to reach is four, three, and two respectively. Therefore the maximum for each of these geodesic diamonds is seven.

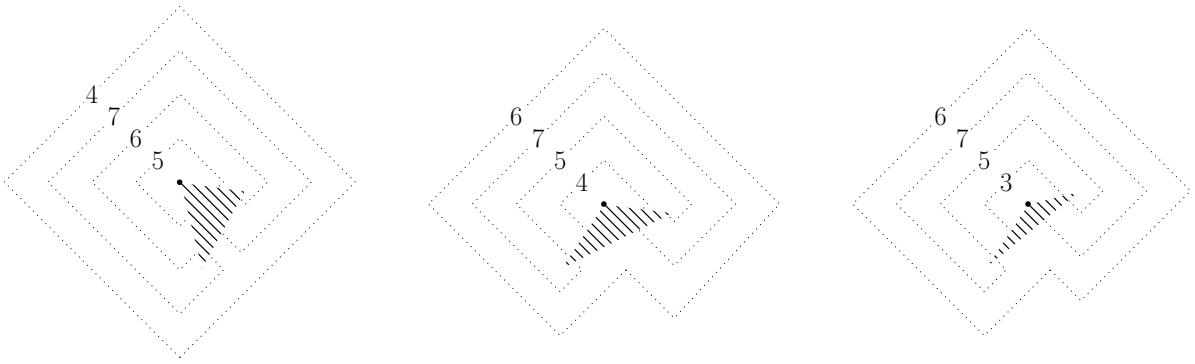


Figure 29: The possible geodesic diamonds centred at corner, edge, or interior vertices of $\mathcal{B}$.

This idea can easily be extended to geodesic diamonds centred on other vertices of $\mathcal{VD}$ since, beyond there being a geodesic diamond of four edges about every point, the possible geodesic diamonds are identical, and the theory does not rely on the vertex being connected to $\mathcal{B}$. □

A.4 Proof of the sufficiency of the partitioning lines

In order to prove that these lines are sufficient we need to note that a bisector will only change if the shortest path from either A_1 or Z in the barrier-less case is altered by the addition of $\mathcal{B}$. We will prove that this only happens, and so the bisector will only change its representation, when Z crosses a line that we have identified above.

It is important initially to realise that a barrier can affect the shortest path only in areas of $\mathcal{P}$ that are not visible from that facility, and that once a barrier has been placed there are now two different possible shortest path routes: one which travels clockwise about $\mathcal{B}$ and the other which travels anticlockwise. Owing to this latter property we can produce different bisectors when the clockwise distance becomes shorter than the anticlockwise distance, and vice versa.

Finally, before we begin our proof, let us recall Lemma 4.1. This result tells us that the bisector will only ever be different to the one in the barrier-free scenario if it enters an area of $\mathcal{P}$ not visible from A_1 or Z. We first argue that the bisector can only ever enter a non-visible part of $\mathcal{P}$ (from A_i, Z, or B_j) if it crosses one of our lines.

Lemma A.1. *If Z is moved so that the bisector $B(A_i, Z)$ in the barrier-free problem moves from outside to inside a non-visible area of $\mathcal{P}$ from either A_1 or Z, then it must have crossed one of our lines.*

Proof. There are two ways in which the bisector can move into such a non-visible area. Either the visible area can change due to the placement of Z, or the change in Z will move the bisector into an area that is non-visible.

Firstly it is clear from Figure 6 that the visible area will only ever change (since A_1 and $\mathcal{B}$ are fixed) if Z crosses a horizontal or vertical line through the extreme vertices of $\mathcal{B}$ (the topmost, leftmost, bottommost, and rightmost). They are exactly quadrant lines through these vertices of $\mathcal{B}$, which we have recorded.

Now it remains to be seen when the bisector will be moved into a non-visible area from a visible area (with the visible area remaining unchanged). We assume that Z does not cross any configuration lines through A_1 (while this could change the bisector orientation to enter a non-visible part, it is already being recorded) so the bisector does not change shape for small changes of Z. Observing the possibilities in Figure 6 for a bisector to enter the non-visible part of $\mathcal{P}$, we see that the bisector will enter a non-visible area either by its intersection with $\mathcal{B}$ passing by an extreme vertex of $\mathcal{B}$ or by the bisector passing by the corner of $Shadow(\mathcal{B})$. This means that it must cross the lines on which the extreme vertex of $\mathcal{B}$ or the corner of $Shadow(\mathcal{B})$ is equidistant between A_1 and Z. These are exactly the geodesic diamonds centred on B_i and the kink lines of the corners of $Shadow(\mathcal{B})$ (note that the kink lines are constrained to the range in which the chosen corner will be non-visible).

Therefore we have found all lines. □

Once a non-visible part is entered, the bisector no longer resembles that from the barrier-free problem. The bisector between A_i and A_j now becomes that of the additive bisector within the non-visible part, where the extreme vertex B_k of $\mathcal{B}$ which is casting the shadow for facility A_i (without loss of generality) now becomes the new facility replacing A_i with an additive value of $l_1(A_i, B_k)$.

Now that the barrier vertices are acting like facilities in the creation of (additive) bisectors, there are new non-visible areas that the bisector could enter, thereby creating additive bisectors of the additive bisector. Therefore we must also check that we are observing when the additive bisector parts are travelling into non-visible parts according to the new bisector points (the vertices of $\mathcal{B}$).

Lemma A.2. *If Z is moved so that the bisector $B^{\mathcal{B}}(A_i, Z)$ moves into a new non-visible area of $\mathcal{P}$, thereby possibly affecting the bisector representation, then it must have crossed one of our lines.*

Proof. By Lemma A.1 we record when the bisector enters a non-visible part from a visible part. Once this happens the bisectors behave as additive bisectors between a facility and a vertex of $\mathcal{B}$. This new bisector could change again once it enters a non-visible area with respect to the facility and the barrier vertex. Therefore we must check that we have recorded whenever a bisector enters a non-visible part of $\mathcal{P}$ according to any facility or any barrier vertex.

However, this is identical to the proof required in Lemma A.1 since from the barrier vertex the visible area is still determined by the extreme vertices of $\mathcal{B}$ because $\mathcal{B}$ is convex. Therefore the bisector can enter a non-visible area from A_1, or the barrier vertex if the bisector passes an extreme vertex of $\mathcal{B}$, or the corner of $Shadow(\mathcal{B})$, thereby crossing the quadrant lines or kink lines. □

Not only do we require the knowledge of when a bisector enters a new non-visible area, thereby affecting the shortest path, but, due to the special quality of l_1 bisectors, once it enters this area we need to know how the bisector presents itself – i.e. how is it configured? The configuration lines in Averbakh et al. (2015) suitably carve up the space in the barrier-free case and this has been extended here to the additive configuration lines which provide the desired partition for each non-visible area into which the additive bisector may venture.

Now that we have proved when the bisector changes route due to the barrier, and how it would then naturally appear, we need to observe when it changes representation due to intersection with the edges of $\mathcal{VD}$. To prove that we have found all of the lines upon which the placement of Z causes the bisector to intersect a vertex of $\mathcal{VD}$ (since placing on either side of this line will result in an intersection on the edges either side of the vertex), we need only observe that, in order for the bisector to intersect a vertex, these lines are exactly the lines of equal distance about the vertex with distance equal to that from the vertex to the facility on the other side of the bisector. These are exactly the geodesic diamonds centred on vertices within $\mathcal{VD}$ to the facility A_k.

At an intersection we must also check what part of the bisector is intersecting these fixed edges, and so we need to know when a breakpoint intersects the edges of $\mathcal{VD}$. As we have learnt above, the bisector $B^{\mathcal{B}}(P, Q)$ will only change direction (create a breakpoint) once it hits either a vertical or horizontal line through P or Q, or a vertical or horizontal line through an extreme barrier vertex if this barrier vertex is acting as an anchor in an additive bisector, or when it enters a non-visible part from P, Q, or the barrier vertex mentioned. Once it enters a non-visible part it will again anchor on a barrier vertex so we need only to examine the breakpoints at P, Q, and extreme vertices of $\mathcal{B}$. Breakpoints occur at the same horizontal or vertical coordinate as the points, and we have these exact lines in the intersection lines.

To prove, if it were in doubt, that these are the only intersections of $\mathcal{B}$ or $\mathcal{P}$, we use Lemma 4.2 which tells us that a barrier can have either zero or two intersections (with touches not counting as an intersection), and it is at these intersections that Z must lie on one of our lines.

This result also brings us to our final point. The intersections divide the bisector into two parts: one part using the distances calculated travelling clockwise around $\mathcal{B}$ and the other part travelling anticlockwise around $\mathcal{B}$. Studying only one of these parts we see that, by Lemma 4.2, if it intersects $\mathcal{B}$ or $\mathcal{P}$ then the other part must also intersect $\mathcal{B}$ or $\mathcal{P}$. However, there is the possibility that these lines meet before intersecting the barrier or boundary at all, which gives us our final possible bisector deformation. This happens only when the clockwise and anticlockwise paths to a point are equidistant, and are equal to that distance from the point to the other facility. Thus we need to discover the lines upon which these bisector parts meet. These, by definition, are the wrap-around lines.

Thus we have identified every way in which the bisector can be changed by the barrier, and proved that we have shown the lines which preserve structural identity in this situation.