Geometric Clusterings[*]

Vasilis Capoyleas [†]
Günter Rote [‡]
Gerhard Woeginger [‡]

submitted to the Journal of Algorithms

revised version, June 1990

Abstract. A k-clustering of a given set of points in the plane is a partition of the points into k subsets ("clusters"). For any fixed k, we can find a k-clustering which minimizes any monotone function of the diameters or the radii of the clusters in polynomial time. The algorithm is based on the fact that any two clusters in an optimal solution can be separated by a line.

AMS 1980 mathematics subject classification (1985 revision): 68Q20, (62H30, 90B99, 52A37)

CR categories and subject descriptors (1987 version): F.2.2. [**Analysis of algorithms and problem complexity**]: Non-numerical algorithms and problems — *geometrical problems and computations*; I.5.3. [**Pattern recognition**]: Clustering — *algorithms*; I.3.5. [**Computer graphics**]: Computational geometry — *geometric algorithms*

General terms: algorithms, theory

1 Introduction

Problem statement. Let S be a set of n points in the plane. A partitioning of S into k disjoint (possibly empty) sets $C_1, C_2, \ldots, C_k$, is called a *k-clustering*, and the individual sets C_i are called its *clusters*. In cluster analysis, the points represent properties (data) of real-world objects, and the aim is to collect "similar" objects (points which are close to each other) in the same cluster, and to put objects which are very "different" into different clusters.

Let W be some weight function that assigns a real weight to any finite set C of points in the plane, like the diameter of C, the radius of C, or the perimeter or the area of its convex hull. Further, let $\mathcal{F}$ be a k-ary symmetric function, assigning a real value to every k-tuple of reals. (Examples for $\mathcal{F}$ are the sum or the maximum.)

The *geometric k-clustering problem for W with respect to $\mathcal{F}$* is defined as follows.

[*]This work was initiated independently by Vasilis Capoyleas and by the other two authors. The work of Günter Rote and Gerhard Woeginger was done while they were at the Freie Universität Berlin, Fachbereich Mathematik, Institut für Informatik; their work was partially supported by the ESPRIT II Basic Research Actions Program of the EC under contract no. 3075 (project ALCOM). They thank the participants of the Mittagsseminar of the theoretical computer science group in Berlin for being such a nice audience and for stimulating discussions. Günter Rote also acknowledges the support by the Fonds zur Förderung der wissenschaftlichen Forschung, Projekt S32/01.

[†]Rutgers University, Computer Science Department, Piscataway, New Jersey 08854, USA

[‡]Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Austria

INSTANCE: A set S of n points in the plane; a rational number d.
QUESTION: Is there a k-clustering for S into k sets $C_1, C_2, \ldots, C_k$ such that $\mathcal{F}(W(C_1), W(C_2), \ldots, W(C_k)) \leq d$?

Previous work and our result. If k is part of the input, this problem is in general NP-complete. Supowit [18] has shown this result for W being the diameter and for $\mathcal{F}$ being the maximum function; in other words, for k part of the input, minimizing the maximum diameter in a k-clustering is NP-complete. The related problem of minimizing the maximum radius, which in the area of location problems is also known as the k-center problem, is also NP-complete (Megiddo and Supowit [14]). It is even NP-hard to find a solution whose maximum radius (or maximum diameter) is within a factor of 1.82 (or 1.97, respectively) of the optimum (Feder and Greene [9]). For fixed k, a polynomial algorithm for minimizing the maximum radius has been given by Drezner [6]. NP-completeness can also be shown for minimizing the maximum cluster area and for minimizing the sum of all cluster areas, as follows from a result of Megiddo and Tamir [15] that it is NP-complete to decide whether a set of points can be covered by a given number of lines. For more information, the interested reader is referred to Brucker [5] and to Johnson's NP-Completeness Column [13].

In this note we show that for every **fixed** k, the geometric k-clustering problem becomes solvable in polynomial time, if W and $\mathcal{F}$ are as follows:

- W is the diameter or the radius;
- $\mathcal{F}$ is an arbitrary monotone increasing function.

Standard examples for monotone increasing functions $\mathcal{F}$ are the *maximum*, the *sum*, or the *sum of the squares* of k non-negative arguments. The 2-clustering problem for the maximum diameter has been treated by Asano, Bhattacharya, Keil, and Yao [1]. They gave an $O(n \log n)$ algorithm for this problem. Monma and Suri [16] gave an $O(n^2)$ algorithm for finding a 2-clustering with smallest sum of diameters.

Some further related results are discussed in the concluding section.

Overview of the paper. The key result that we will use is that for any given 2-clustering, there is always a 2-clustering which is at least as good as the given one (as regards the diameters or the radii of *both* clusters in each clustering) and in which the two clusters can be separated by a line. For the case of radii, this is easy to see, whereas the proof for the case of diameters is more elaborate. It is the subject of section 2. This theorem allows us to limit the number of possible candidates for optimal solutions. From this, a polynomial-time algorithm, which essentially tests all these candidates, follows in a quite straightforward way. This algorithm is derived in section 3.

The problem of testing whether a 2-clustering with specified bounds on the two diameters exists has been treated by Hershberger and Suri [11]. They gave an $O(n \log n)$-time algorithm which does not use the separability of the two clusters.

Our paper establishes the polynomial complexity status of a class of clustering problems, and it gives some insight into the structure of optimal solutions. Our results hold in a quite general setting. We are of course far from claiming that the algorithms that follow from our paper are optimal, for any specific problem.

1.1 Definitions, Notations, Elementary Facts

We now give some definitions and notations and we state some elementary geometric facts that will be used in the paper. The convex hull of a point set A is denoted by $conv(A)$, the *diameter* (the maximum distance of two points in A) by $diam(A)$. By the *perimeter* of a point set A we mean the perimeter of it convex hull, i. e., the length of the boundary of $conv(A)$. The *radius* $r(A)$ of a finite point set A is the radius of the smallest enclosing circle. We can define the radius, the perimeter, and the diameter of the empty set as 0 or $-\infty$, as we like. Two sets are said to be *linearly separable*, if they can be strictly separated by a straight line. It is well known that two sets are linearly separable if and only if their convex hulls are disjoint. The Euclidean distance of two points p_1 and p_2 is denoted by $d(p_1, p_2)$.

Proposition 1

(i) *In a convex quadrangle* $\square abcd$, *the sum of the lengths of the diagonals is always larger than the sum of two opposite sides:* $d(a,c) + d(b,d) > d(a,b) + d(c,d)$ *and* $d(a,c) + d(b,d) > d(a,d) + d(b,c)$.

(ii) *In a triangle with an obtuse angle, the side lying opposite the obtuse angle is the longest side in the triangle.*

(iii) *Let* d *be a positive real, let* p_1 *and* p_2 *be two points in the plane at distance less than or equal to* d*. Let* C_1 *and* C_2 *be the circles with radius* d *centered at* p_1 *and* p_2*, let* D *denote the points in the vertical stripe between* p_1 *and* p_2*. Then the part of the region* $C_1 \cap C_2 \cap D$ *that lies above the line through* p_1 *and* p_2 *has diameter* $\leq d$ *(cf. Figure 7).*

Proof. We only prove (iii). Recall that the diameter of a convex figure is equal to the greatest distance between two parallel supporting lines of the figure (see Preparata and Shamos [17, Section 4.2.3]). One of two parallel supporting lines must touch the figure at p_1 or p_2. As all points in the figure lie within the circles C_1 and C_2, the other supporting line is at distance at most d. $\square$

2 Separability of Two Clusters in the Diameter Case

Our results about the case where the function W is the diameter are based on the following theorem which shows how we can separate two intersecting clusters by a line without increasing the diameters. This will imply that we can assume w. l. o. g. that all clusters in an optimal solution are pairwise separable.

Theorem 2 *Let* A *and* B *be two sets of points in the plane with diameters* d_A *and* d_B*. Then there are two linearly separable sets* A' *and* B' *with diameters* $d_{A'}$ *and* $d_{B'}$ *such that* $d_{A'} \leq d_A$, $d_{B'} \leq d_B$ *and* $A' \cup B' = A \cup B$.

Proof. The proof will proceed in several elementary steps and involve two intermediate lemmas.

We may assume that $conv(A) \cap conv(B) \neq \emptyset$, $conv(A) \not\subseteq conv(B)$, and $conv(B) \not\subseteq conv(A)$, as otherwise the statement is trivial. If $conv(A) \cap conv(B)$ consists only of a line segment or a single point, we either set $A' = (A \cup B) \cap conv(A)$ and $B' = (A \cup B) - conv(A)$, or we set $A' = (A \cup B) - conv(B)$ and $B' = (A \cup B) \cap conv(B)$. It is straightforward to check that in at least one of these two cases the convex hulls $conv(A')$ and $conv(B')$ are disjoint, and hence we are done.

Otherwise, let $\langle u_1, u_2, \ldots, u_{2k} \rangle$ be the sequence of points where the boundaries of *conv*(A) and *conv*(B) cross, in clockwise order. We write *conv*$(A) -$ *conv*(B) and *conv*$(B) -$ *conv*(A) as two interlacing sequences of polygons $\langle A_1, A_2, \ldots, A_k \rangle$ and $\langle B_1, B_2, \ldots, B_k \rangle$, where A_i touches B_i at u_{2i}, and A_i touches B_{i-1} at u_{2i-1} (see Figure 1). The polygons A_i and B_i have the shape of half-moons: Their boundary consists of a convex and a concave chain. (The boundaries of *conv*(A) and *conv*(B) may touch without crossing each other, or they may even coincide on a small piece. Thus, a half-moon A_i or B_i may have degenerate parts where its boundary "touches itself" or parts which are line segments, and the points u_j may not be defined uniquely. This, however, will not affect our arguments.)

We will separate the set $A \cup B$ into A' and B' by a line L through two points u_i and u_j, whose choice will be described.

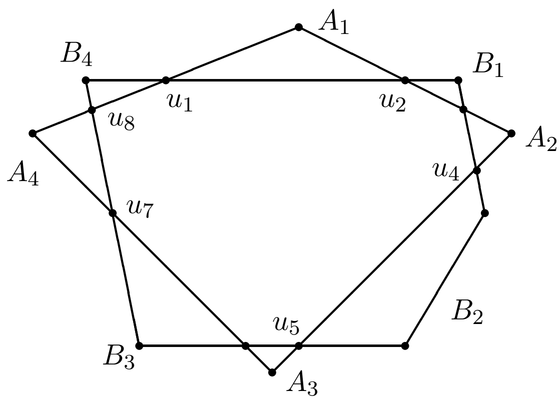

Figure 1: Regions created by two intersecting convex polygons.

Without loss of generality, we may assume that $d_A \geq d_B$. We call a pair (A_i, B_j) a *bad pair*, if $diam(A_i \cup B_j) > d_A$. B_j is called a *bad partner* of A_i, and vice versa. The bad pairs are those pairs of half-moons which must be separated by the line L in order to make both diameters $\leq d_A$. A half-moon A_i or B_j is called a *bad half-moon*, if it appears in some bad pair. First, we will prove an intermediate lemma about the relative positions of two bad pairs. We say that two pairs (A_i, B_j) and $(A_{i'}, B_{j'})$ with $A_i \neq A_{i'}$ and $B_j \neq B_{j'}$ *cross* if their cyclic sequence is $A_i, A_{i'}, B_j, B_{j'}$ or $A_i, B_{j'}, B_j, A_{i'}$. In other words, they cross if and only if the two segments connecting a point in A_i to a point in B_j and a point in $A_{i'}$ to a point in $B_{j'}$ intersect, independent of the choice of these points. Such segments are called *bad* segments.

Lemma 3 *Any two disjoint bad pairs cross.*

Proof. Let us assume that there are two bad pairs (A_i, B_j) and $(A_{i'}, B_{j'})$ with $A_i \neq A_{i'}$ and $B_j \neq B_{j'}$ that do not cross. For each bad pair, we choose a bad line segment connecting two points at distance $> d_A$ that lie in the bad half-moons belonging to the pair. Let us call these points a_i, b_j, $a_{i'}$, and $b_{j'}$, respectively. The two possibilities for the relative positions of these points (disregarding symmetric variations) are depicted in Figure 2. The bad segments are represented by double lines, their endpoints are shown as black circles (points in A) and white circles (points in B).

(a) The case shown on the left side of Figure 2 immediately leads to a contradiction: By

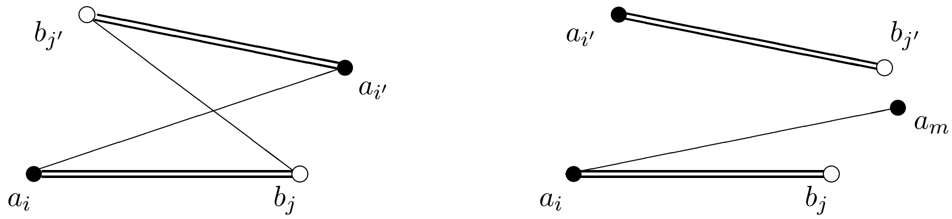

Figure 2: Two impossible configurations of bad pairs.

Proposition 1.i, the sum of the diagonals in the convex quadrangle $\square a_i b_j a_{i'} b_{j'}$ is larger than the sum of two opposite sides. Hence,

$$d_A + d_B \geq d(a_i, a_{i'}) + d(b_j, b_{j'}) > d(a_i, b_j) + d(b_{j'}, a_{i'}) > 2d_A$$

must hold, a contradiction to the assumption $d_A \geq d_B$.

(b) In the case shown on the right side of Figure 2, we observe that the convex quadrangle $\square a_i b_j b_{j'} a_{i'}$ must have an angle larger or equal to $\pi/2$. W. l. o. g., let this be the angle at b_j. Between the half-moons B_j and $B_{j'}$, there lies at least one (not necessarily bad) half-moon A_m. Select an arbitrary point $a_m \in A_m$. Then the angle $\angle a_i b_j a_m$ is obtuse, and hence, by Proposition 1.ii,

$$d(a_i, a_m) > d(a_i, b_j) > d_A.$$

This is again a contradiction. $\square$

For stating our next lemma, we group adjacent bad half-moons from the same cluster (A or B) together. Thus, we define a *group of bad half-moons* to be a maximal cyclic subsequence of bad half-moons from one cluster. (Intervening half-moons of the other cluster must not be bad.)

Lemma 4 *All bad partners of the half-moons in a group belong to the same group.*

Proof. Assume that two half-moons A_i and $A_{i'}$ belonging to the same group form bad pairs (A_i, B_j) and $(A_{i'}, B_{j'})$ with two half-moons B_j and $B_{j'}$ which are in different groups. (A_i and $A_{i'}$ may be the same). Then there must be a bad half-moon $A_{i''}$ between B_j and $B_{j'}$. But this half-moon cannot have a bad partner $B_{j''}$ without forming two disjoint non-crossing bad pairs with either (A_i, B_j) or $(A_{i'}, B_{j'})$, contradicting Lemma 3. (cf. figure 3). $\square$

Now we know that the bad pairs give rise to a complete matching among the groups. Since bad pairs must cross, there is an odd number of groups from each cluster, and they must be completely interlacing, as shown in figure 4.

Now it will be easy to achieve our first goal in proving Theorem 2 — finding a line that separates all bad pairs: As discussed previously, any such line ensures that $d_{A'} \leq d_A$ and $d_{B'} \leq d_A$. In addition, we also want to obtain the inequality $d_{B'} \leq d_B$. Therefore, among the possible separating lines which cut all bad pairs we will select one which makes the smaller part as small as possible, i. e., which cuts as unbalanced as possible.

In the rest of this section we will only discuss the case that there is more than one (i. e., at least three) group of bad half-moons of each cluster, as the other cases are analogous but simpler.

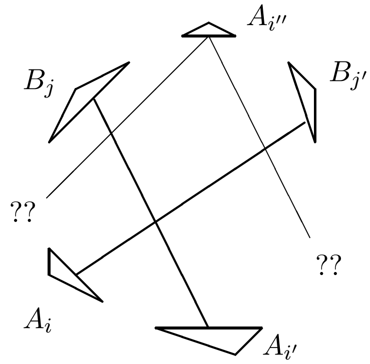

Figure 3: No place for bad partners of $A_{i''}$.

[HAND-DRAWN FIGURE]

Figure 4: The structure of bad pairs between groups of bad half-moons.

We construct our line L as follows: Let A_i be the last bad half-moon of a group (in clockwise order), and let $B_{j'}$ be the last bad partner of A_j. Let B_j be the first bad half-moon after A_i, and let $A_{i'}$ be the first bad partner of B_j (see Figure 5). We choose the separating line L to go through the point u_{2j} before B_j and the point $u_{2j'+1}$ after $B_{j'}$. We define B' to be the points in $A \cup B$ lying on the same side of L as B_j and $B_{j'}$, and A' as the remaining points.

As discussed above, L cuts all bad pairs, and both diameters $d_{A'}$ and $d_{B'}$ are $\leq d_A$. It remains to show that $d_{B'} \leq d_B$ holds.

Let us first make a few observations about the half-moons A_i, B_j, $B_{j'}$, and $A_{i'}$. Since B_j and $B_{j'}$ belong to different groups, (A_i, B_j) is no bad pair, and similarly, $(A_{i'}, B_{j'})$ is no bad pair. Moreover, the clockwise sequence of the four half-moons is $A_i, B_j, B_{j'}, A_{i'}$, and they are all different. As A_i and B_j were neighbors in the clockwise ordered sequence, no bad half-moon lies between A_i and B_j, and similarly, no bad half-moon lies between $B_{j'}$ and $A_{i'}$. We can select a_i, b_j, $b_{j'}$, and $a_{i'}$ in the four half-moons such that $d(a_i, b_{j'}) > d_A$ and $d(a_{i'}, b_j) > d_A$.

Let L be drawn horizontally, with B' above L, as in Figure 5. Consider two points $a, b \in B'$. We have to show that $d(a, b) \leq d_B$. If $a, b \in B$ then there is nothing to prove.

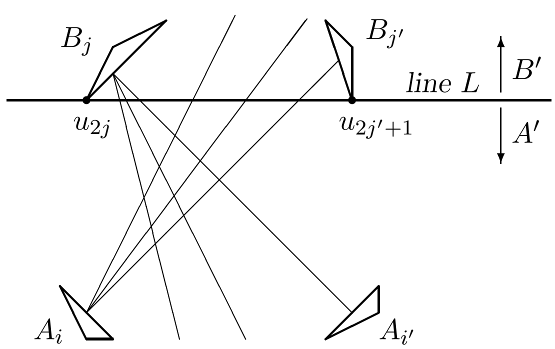

Figure 5: How to find the separating line.

Otherwise, we proceed in three steps: We first consider the case $a \in A - B$, $b \in B - A$, and then the case $a \in A - B$, $b \in A - B$, and finally the case $a \in A - B$, $b \in A \cap B$.

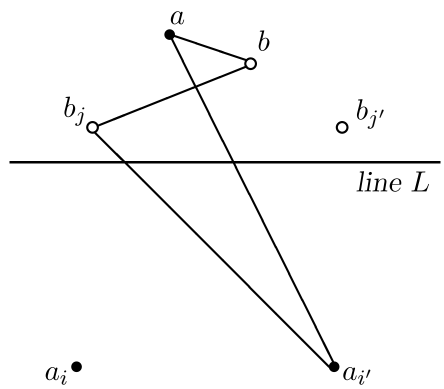

Figure 6: An impossible configuration.

(a) Assume that there were two points $a \in A_k$ and $b \in B_l$ in two (not necessarily bad) half-moons belonging to B' such that $d(a,b) > d_B$ holds. Consider first the case that a comes before b in the clockwise ordering, see Figure 6. By Proposition 1.i, this would imply

$$d_A + d_B \geq d(a, a_{i'}) + d(b, b_j) > d(b_j, a_{i'}) + d(a, b) > d_A + d_B.$$

In the other case, when b comes before a, we have symmetrically

$$d_A + d_B \geq d(a, a_i) + d(b, b_{j'}) > d(a_i, b_{j'}) + d(a, b) > d_A + d_B.$$

(b) We now consider the possibility that two points in $(A-B) \cap B'$ might be more than d_B apart. Assume, some point $a \in A_k$ in a half-moon belonging to B' lies to the left of b_j. Then the angle $\angle ab_ja_{i'}$ is obtuse, and from Proposition 1.ii we get the contradiction

$$d(a, a_{i'}) > d(b_j, a_{i'}) > d_A.$$

The same holds, if a lies to the right of $b_{j'}$. From (a), applied to the point a together with b_j and $b_{j'}$, resp., we get that all points a in $(A-B) \cap B'$ lie at distance $\leq d_B$ from b_j and from $b_{j'}$. Moreover, they are all above the line through b_j and $b_{j'}$ (see Figure 7). Hence, from Proposition 1.iii it follows that $(A - B) \cap B'$ has diameter $\leq d_B$.

(c) Now assume there are two points $a \in A_k$ and $b \in A \cap B$ such that $d(a,b) > d_B$. We may assume that b is a vertex of $conv(A) \cap conv(B) \cap L^+$, where L^+ denotes the halfplane above L, since the distance from a is maximized at some vertex. The vertices of this set belong to the boundary of some B_l or A_l, and hence we can apply the analysis of case (a) or case (b), resp.

Therefore, the diameter of B' is $\leq d_B$, and the proof of Theorem 2 is complete. □

Remark. A weaker version of Theorem 2 was stated in [1]. The proof relied on the erroneous assumption that all bad half-moons A_i can be separated from all bad half-moons B_j by a straight line. The set of six points in Figure 8 gives a counter-example to this assumption. (The triangles $\Delta a_1a_2a_3$ and $\Delta b_1b_2b_3$ are equilateral and equal, and all six points lie on a common circle.)

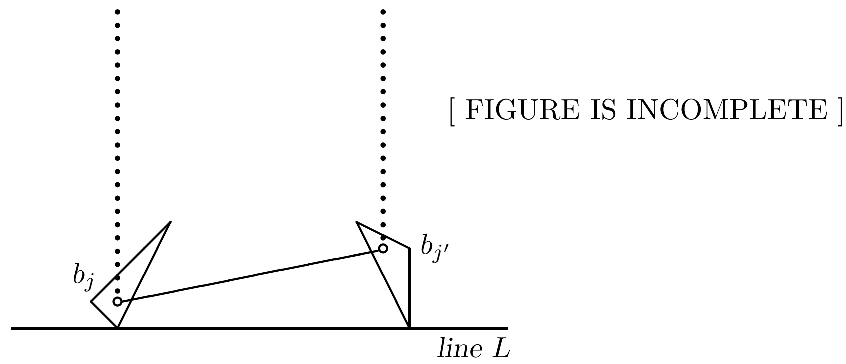

Figure 7: $(A - B) \cap B'$ is a subset of the shaded region.

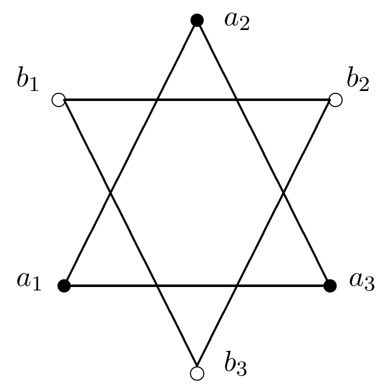

Figure 8: A counter-example.

Lemma 5 *In the construction in Theorem 2,*

$$perimeter(A) + perimeter(B) \geq perimeter(A') + perimeter(B')$$

holds. If $conv(A) \cap conv(B) \neq \emptyset$, then the inequality is strict.

Proof. If $conv(A) \cap conv(B) = \emptyset$, we have $A' = A$ and $B' = B$, and there is nothing to prove. In the degenerate case where $conv(A) \cap conv(B)$ consists only of a line segment or a single point, we have $conv(A') = conv(A)$ and $conv(B') \subsetneq conv(B)$, or vice versa, and the claim holds again. In the remaining case, where $conv(A) \cap conv(B)$ has non-empty interior, we have to make a small calculation. We use the following notations (for an illustration, see Figure 9): By *Length($\cdots$)*, we denote the perimeter of $conv(A) \cap conv(B)$, the length of the dotted lines in the figure. By *Length(—)*, we denote $perimeter(A) + perimeter(B) - Length(\cdots)$, the length of the boundary of $conv(A) \cup conv(B)$ shown by the solid line in the figure. By *Length(//)*, we denote twice the length of the line segment that results from intersecting the separation line L with $conv(A)$, shown as a double line segment in the figure. Now it is easy to see that

$$\begin{aligned} perimeter(A') + perimeter(B') &\leq Length(—) + Length(//) \\ &< Length(—) + Length(\cdots) \\ &= perimeter(A) + perimeter(B) \end{aligned}$$

holds, since two curves of total length $\mathit{Length}(—) + \mathit{Length}(/\!\!/)$ enclose the two convex sets $conv(A)$ and $conv(B)$. The second inequality is strict because $conv(A) \cap conv(B)$ has non-empty interior. □

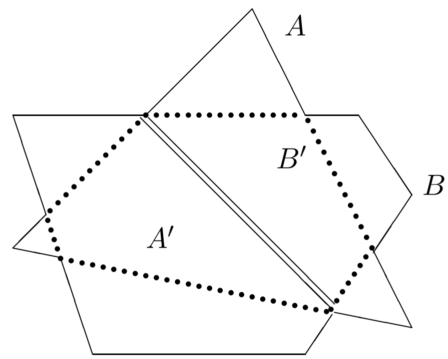

Figure 9: How to subdivide the perimeters.

3 The Polynomial Time Result

In this section, we extend Theorem 2 of the preceding section to more than two clusters, and we also show the corresponding result for the case of radii. Finally, we will apply these separability theorems to obtain a polynomial algorithm.

Theorem 6 *Consider the optimal k-clustering problem for the diameter with a monotone increasing function $\mathcal{F}$. For every point set P in the plane, there is an optimal k-clustering such that each pair of clusters is linearly separable.*

Proof. Consider the optimal k-clustering for which the sum of the perimeters of all clusters becomes minimal. Assume that there are two clusters which are not linearly separable. Applying Theorem 2 and Lemma 5 to the two clusters, we get a k-clustering with smaller sum of perimeters. As both affected diameters do not increase, the value of $\mathcal{F}$ does not increase, too. □

So far, we have only dealt with the diameter as the quality measure of a cluster. For the radius, an analog of Theorem 6 can be shown directly.

Theorem 7 *Consider the optimal k-clustering problem for the radius with a monotone increasing function $\mathcal{F}$. For every point set P in the plane, there is an optimal k-clustering such that each pair of clusters is linearly separable.*

Proof. Let a k-clustering with radii $r(C_1), \ldots, r(C_k)$ be given. We shall show that there is a decomposition of the plane into k convex polygonal cells $R_1, \ldots, R_k$ such that the clustering $R_1 \cap P \ldots, R_k \cap P$ corresponding to this decomposition is at least as good as the given one, i. e., $r(R_i \cap P) \leq r(C_i)$, for all $i = 1, \ldots, k$.

We know that each cluster C_i is contained in a disk D_i with radius $r_i = r(C_i)$ and center M_i. Suppose that two disks D_1 and D_2 intersect. The most natural choice of a line separating the two clusters is the line L through the intersection points of the two circles. If we reassign points in $D_1 \cap D_2$ to the two clusters according to their position relative to L, we see that the new clusters are still contained in their respective disks, and thus the new cluster radii cannot become larger.

We would like to perform this reassignment of points for each pair of clusters. The line L consists of the points x fulfilling the following equation

$$d(x, M_1)^2 - r_1^2 = d(x, M_2)^2 - r_2^2.$$

The expression $d(x, M_1)^2 - r_1^2$ is called the *power* of the point x with respect to the disk D_1. Its sign indicates whether x is contained in D_1. L is called the power line, radical axis, or chordale of the two disks. It is defined for any pair of non-concentric disks. If we partition the plane by assigning every point to the disk D_i for which its power $d(x, M_i)^2 - r_i^2$ is minimal, we get the so-called *power diagram* (cf. Aurenhammer [2]; Imai, Iri, and Murota [12]; or Edelsbrunner [7], section 13.6). It is known that the power diagram is a dissection of the plane into (at most) k convex polygonal regions, very much like in the case of Voronoi diagrams, which are a special case of power diagrams where all radii are equal.

Since for each point $p \in P$, $d(x, M_i)^2 - r_i^2 \leq 0$ for at least one i, the power is also ≤ 0 for the disk belonging to the region to which p is assigned; in other words, p is contained in this disk. Thus, the power diagram is the desired dissection. □

Since a planar dissection into k convex polygonal regions has at most $3k - 6$ edges (for $k \geq 3$), we will only have to specify a straight line for each of these edges in order to get a possible candidate for an optimal clustering. Since the number of such choices is limited, this will give our polynomial-time result.

We remark that, for the diameter, such a decomposition into convex regions need not necessarily exist. We know that an optimal solution is specified by $\binom{k}{2}$ lines, one for each pair of vertices. This is much larger than $3k - 6$. However, the following reformulation of a result of Edelsbrunner, Robison, and Shen [8, Lemmas 1 and 2] shows a somewhat weaker statement than the above theorem, which is nevertheless completely sufficient for our purposes.

Lemma 8 *Let $R_1, \ldots, R_k$ be k convex, compact, and pairwise disjoint sets in the plane. Then there is a planar graph $G = (V, E)$ whose vertices are the k given sets, with the following properties:*

- *For each edge $\{R_i, R_j\} \in E$, there is a line which cuts the plane into two open half-planes H_{ij} and H_{ji}, such that R_i is contained in H_{ij} and R_j is contained in H_{ji}; i. e., this line strictly separates R_i from R_j.*
- *For each R_i:*
$$R_i \subseteq R_i' := \bigcap_{\{R_i, R_j\} \in E} H_{ij}. \qquad (*)$$
- *The regions R_i' are disjoint.* □

In the proof of this statement one lets the regions R_i grow until they are maximal non-overlapping convex cells. The planar graph G of the lemma is a kind of dual graph of the resulting polygonal sets, having an edge for every pair of touching regions.

Now we can state our main result:

Theorem 9 *For any fixed k, the geometric k-clustering problem for the diameter or for the radius with respect to some monotone increasing function $\mathcal{F}$ is solvable in $O(n^{6k})$ time.*

Proof. The convex hulls $R_i = conv(C_i)$ of the clusters in a k-clustering are convex and compact. By Theorem 6 or Theorem 7, resp., we may assume that they are pairwise disjoint. Lemma 8 shows then that we can completely determine a solution as follows:

1. We have to choose a planar graph $G = (V, E)$ with k vertices.
2. For each edge $\{i, j\} \in E$, we have to select a line and we have to specify which side of this line is to contain C_i and which side should contain C_j.
3. Then, we determine for each point $p \in P$ to which sets H_{ij} it belongs, and we evaluate $(*)$. If each point happens to fall into exactly one cluster, we have a possible candidate for an optimal solution.

Let us estimate the time to generate all possible candidates: There is only a fixed number of non-isomorphic planar graphs with k vertices. The number of edges is at most $3k - 6$, for $k \geq 3$. The number of different ways in which a set of n points can be separated by a line into two sets (P_1, P_2) is $n(n-1)+2$. Thus, the number of possibilities of steps 1 and 2 is $(n(n-1)+2)^{3k-6} = O(n^{6k-12})$. The check in step 3 takes $O(n)$ time (for fixed k). Finally, if we have found a clustering, we determine the diameters or the radii in $O(n \log n)$ time or in $O(n)$ time, resp., (see Preparata and Shamos [17, Sections 4.2.3 and 7.2.5]), and we evaluate $\mathcal{F}$. Clearly, the minimum over all values that we get is the solution to the k-clustering problem. Thus, assuming that the evaluation of $\mathcal{F}$ takes reasonable time (not more than, say $O(k^{12})$), we get the desired result. $\square$

Remark: In the case of radii, we would not need Lemma 8. As the graph G, we can take the dual of the power diagram in the proof of Theorem 7.

4 Concluding Remarks

(1) A related separability result was known for the problem of minimizing the sum of the variances of the clusters. The variance of a cluster is the sum of the squares of the distances of all pairs of points in the cluster, divided by the number of points (cf. Bock [4, section 15, pp. 162–176]).

$$\mathrm{Var}(A) := \sum_{\{a_i, a_j\} \subseteq A} d(a_i, a_j)^2 / |A| = \sum_{a_i \in A} d(a_i, \bar{a})^2,$$

where $\bar{a} = \sum_{a_i \in A} a_i / |A|$ is the center of gravity of the cluster. Using the above expression for the variance, it is straightforward to show that two clusters of an optimal clustering are always separated by the bisecting line (or hyperplane) of the cluster centers. In fact, an optimal clustering is induced by the Voronoi diagram of its cluster centers.

Similarly, for the problem where the sum of the squares of all distances between points in the same cluster is to be minimized (without division by the cluster sizes), Boros and Hammer [3] showed that two clusters in an optimal solution can always be separated by circle (or a sphere, in higher dimensions).

In both of these cases the separability result is due to the special form of the objective function.

(2) Let us discuss how we would go about actually finding a 3-clustering which minimizes the *maximum* diameter, using the results at hand. We know that one of the clusters is separable from the rest of the points by two lines. If we take this cluster away (there are at most $(n(n-1)+2)^2 = O(n^4)$ ways to do this), we can solve the remaining 2-clustering problem in $O(n \log n)$ time by the optimal algorithm of Asano, Bhattacharya, Keil, and Yao [1]. This yields a complexity of $O(n^5 \log n)$ for 3-clusterings. Similarly, for 4-clusterings, we can separate the problem into two 2-clusterings by 4 lines, yielding a complexity of $O(n^9 \log n)$.

The above examples show that, already for $k = 3$ or $k = 4$, there is still ample space for improvements.

(3) In contrast to Theorem 7 (separability of clusters when a function of the radii is minimized), the corresponding theorems for the diameter case (Theorem 2 and Theorem 6) do not generalize to higher dimensions, as is shown by the following example: Consider the point set $\{A, B, C, A', B', C'\}$ in three-dimensional space, where $A = (-1, 0, 0)$, $B = (1, 0, 0)$, $C = (0, \sqrt{3}, 0)$, $A' = (-1, 0, -\varepsilon)$, $B' = (1, 0, -\varepsilon)$ and $C' = (0, \varepsilon, \sqrt{2} + \varepsilon)$, for some small $\varepsilon < 1/100$. If the maximum cluster diameter is to be minimized, this point set has only one optimal 2-clustering, $C_1 = \{A, B, C\}$ and $C_2 = \{A', B', C'\}$, having maximum diameter 2. This 2-clustering is not linearly separable.

(4) The proof of Lemma 5 shows that separability of the clusters also holds when we minimize the *sum of the perimeters* of the clusters, and hence our algorithm applies. We do not know whether such a result holds for all monotone functions $\mathcal{F}$ for the case of perimeters.

(5) When we take the area of the convex hull as the quality measure, two clusters need not be separable. This can be shown by simple examples where the points (almost) lie on two lines. This is an indication that the area is not a good measure for the quality of a clustering.

References

[1] Tetsuo Asano, B. Bhattacharya, M. Keil, and F. Yao, Clustering algorithms based on minimum and maximum spanning trees, in: *Proc. Fourth Ann. Symp. Computational Geometry*, Urbana-Champaign, June 1988, Association for Computing Machinery 1988, pp. 252–257.

[2] F. Aurenhammer, Power diagrams: properties, algorithms, and applications, *SIAM J. Computing* **16** (1987), 78–96.

[3] E. Boros and P. L. Hammer, On clustering problems with connected optima in Euclidean spaces, *Discrete Mathematics* **75** (1989), 81–88.

[4] H. H. Bock, *Automatische Klassifikation*, Vandenhoeck & Ruprecht, Göttingen 1974.

[5] P. Brucker, On the complexity of clustering problems, in: R. Henn, B. Korte, and W. Oettli, eds., *Optimization and Operations Research*, Springer Verlag 1977, pp. 45–54.

[6] Z. Drezner, The *p*-centre problem — heuristic and optimal algorithms, *J. Operational Research Society* **35** (1984), 741–748.

[7] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer Verlag, 1987.

[8] H. Edelsbrunner, A. D. Robison, and X. Shen, Covering convex sets with non-overlapping polygons, *Discrete Math.* **81** (1990), 153–164.

[9] T. Feder, D. H. Greene, Optimal algorithms for approximate clustering, in: *Proc. 20th Ann. ACM Symp. Theory Computing*, Association for Computing Machinery 1988, pp. 434–444.

[10] L. Fejes Tóth, Some packing and covering theorems, *Acta Sci. Math. Szeged* **12A** (1950), 62–67.

[11] J. Hershberger and S. Suri, Finding tailored partitions, in: *Proc. Fifth Ann. Symp. Computational Geometry*, Saarbrücken, West Germany, June 1989, Association for Computing Machinery 1989, pp. 255–265; also in *J. Algorithms* **12** (1991), 431–463.

[12] H. Imai, M. Iri, and K. Murota, Voronoi diagrams in the Laguerre metric and its applications, *SIAM J. Computing* **14** (1985), 93–105.

[13] D. S. Johnson, The NP-completeness column: An ongoing guide, *Journal of Algorithms* **3** (1982), 182–195.

[14] N. Megiddo and K. J. Supowit, On the complexity of some common geometric location problems, *SIAM J. Computing* **13** (1984), 182–196.

[15] N. Megiddo and A. Tamir, On the complexity of locating linear facilities in the plane, *Operations Research Letters* **1** (1982), 194–197.

[16] C. Monma and S. Suri, Partitioning points and graphs to minimize the maximum or the sum of diameters, in: *Graph Theory, Combinatorics and Applications*, Vol. 2, Proc. Sixth Quadrennial Int. Conf. Theory and Appl. of Graphs, Kalamazoo, Michigan, May 1988, ed. Y. Alavi et al.; Wiley, 1991, pp. 899–912.

[17] F. P. Preparata and M. I. Shamos, *Computational Geometry — an Introduction*, Springer Verlag, 1985.

[18] K. J. Supowit, Topics in Computational Geometry, Ph. D. Thesis, 1981, Dept. of Computer Science, University of Illinois at Urbana-Champaign, Report UIUCDCS-R-81-1062.

Notes added in proof:

By using Drezner's ideas [6], the geometric k-clustering problem for the radius with respect to an arbitrary monotone and symmetric function $\mathcal{F}$ can be solved by checking only $O(n^{3k})$ possibilitites, because the radius and the covering circle of each cluster are determined by at most three points. This improves the bound of Theorem 9 for the case of radii and makes the separability result of Theorem 7 uninteresting from the point of view of its algorithmic implications.

The same construction as in the proof of Lemma 8 by Edelsbrunner et al. [8] (cf. the brief remarks after Lemma 8) was already used by Fejes Tóth [10] in the theory of packing.