

## On the Complexity of Polyhedral Separability

Nimrod Megiddo

IBM Almaden Research Center, San Jose, CA 95120-6099, USA, and  
Tel Aviv University, Tel Aviv, Israel

**Abstract.** It is NP-complete to recognize whether two sets of points in general space can be separated by two hyperplanes. It is NP-complete to recognize whether two sets of points in the plane can be separated with  $k$  lines. For every fixed  $k$  in any fixed dimension, it takes polynomial time to recognize whether two sets of points can be separated with  $k$  hyperplanes.

### 1. Introduction

The following problem is well known in computational geometry.

**Problem (Linear Separability).** Given two sets of points with integer coordinates  $P = \{\pi^1, \dots, \pi^n\} \subset R^d$  and  $Q = \{\rho^1, \dots, \rho^q\} \subset R^d$ , recognize whether there exists a hyperplane  $H = \{z \in R^d: z^T x = x_0\}$  (characterized by a nonzero vector  $x \in R^d$  and a scalar  $x_0$ ) that separates the sets  $P$  and  $Q$  in the sense that for each point  $\pi^i \in P$ ,  $\pi^i{}^T x < x_0$  and for each point  $\rho^i \in Q$ ,  $\rho^i{}^T x > x_0$ .

It is obvious that the linear separability problem can be formulated as a linear programming problem, and hence is solvable in polynomial time. When two sets cannot be separated by a hyperplane, a natural problem is to find the minimum number of hyperplanes that is required for the separation. Very interesting results were recently obtained by Edelsbrunner and Preparata [3] for the convex two-dimensional case, that is, for the problem of separating two sets of points in the plane by a convex polygon with a minimum number of edges. Also, [1] deals with the problem of separating two nested convex polygons by a polygon with a minimum number of edges. The problem of separating two sets of points in  $R^2$  by a circle was first considered in [4]. It turns out that separability of two sets of points in  $R^d$  by a sphere can be decided in linear time if  $d$  is fixed [8], using methods like the one in [5] and [6].

It is interesting to examine the generalization of linear separability where one is interested in separating two sets of points with  $k$  hyperplanes rather than one. We discuss the general case later but first we consider the case  $k = 2$ .

**Problem (2-Linear Separability).** Given two sets of points with integer coordinates  $P = \{\pi^1, \dots, \pi^p\} \subset R^d$  and  $Q = \{\rho^1, \dots, \rho^q\} \subset R^d$ , recognize whether there exist two hyperplanes  $H_1 = \{z: x^T z = x_0\}$  and  $H_2 = \{z: y^T z = y_0\}$  ( $x, y \in R^d, x_0, y_0 \in R$ ) that separate the sets  $P$  and  $Q$  in the sense expressed by the following conditions:

- (i) For each point  $\pi^i \in P$ , both  $(\pi^i)^T x < x_0$  and  $(\pi^i)^T y < y_0$ .
- (ii) For each point  $\rho^j \in Q$ , either  $(\rho^j)^T x > x_0$  or  $(\rho^j)^T y > y_0$ .

Note that in the definition of 2-linear separability the sets  $P$  and  $Q$  do not play symmetric roles. Surprisingly, the 2-linear separability problem is NP-complete as we show in Section 2. Next we consider the general case of  $k$  hyperplanes (a precise definition is given in Section 3). We show that already in  $R^2$  the problem of separating two sets of points by  $k$  lines (not necessarily forming a convex polygon) is NP-complete. We then argue that separability with any fixed number of hyperplanes (not necessarily forming a convex polyhedron) in any fixed dimension can be decided in polynomial time.

## 2. Separability with Two Hyperplanes

It is easy to show that if two sets of points in  $R^d$  with integer coordinates are separable by two hyperplanes then there exist such separating hyperplanes with rational coefficients, so that the size of the binary representation of the hyperplanes is bounded by a polynomial in the size of the binary representation of the points. Thus it is obvious that 2-linear separability is in NP.

To prove the NP-completeness of 2-linear separability, we first introduce a subclass of satisfiability problems, which we call *reversible satisfiability*.

**Definition.** Let  $\varphi$  be a boolean formula and let  $\bar{\varphi}$  denote the formula obtained from  $\varphi$  by negating each variable. For example, if  $\varphi = (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee \bar{z})$  then  $\bar{\varphi} = (\bar{x} \vee y \vee \bar{z}) \wedge (x \vee z)$ . The formula  $\varphi$  is called *reversible* if either both  $\varphi$  and  $\bar{\varphi}$  are satisfiable or both are not satisfiable.

An obvious example of a reversible formula is  $(x \vee y \vee z) \wedge (\bar{x} \vee \bar{y} \vee \bar{z})$ . The *reversible satisfiability problem* is to recognize the satisfiability of reversible formulae in conjunctive normal form (CNF). It is easy to prove that the reversible satisfiability problem is NP-complete.

**Proposition 1.** *The reversible satisfiability problem with six literals per clause is NP-complete.*

*Proof.* We show that the regular satisfiability problem for CNF formulae with three literals per clause is reducible to the satisfiability problem of reversible CNF formulae with six literals per clause (“reversible 6-SAT”). The reduction goes as follows. For each variable  $x$ , replace each occurrence of  $x$  by  $(x_1 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_2)$ . Similarly, replace each occurrence of  $\bar{x}$  by  $(\bar{x}_1 \vee x_2) \wedge (x_1 \vee \bar{x}_2)$ . Thus, a clause of the form  $x \vee \bar{y} \vee z$  is replaced by

$$[(x_1 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_2)] \vee [(\bar{y}_1 \vee y_2) \wedge (y_1 \vee \bar{y}_2)] \vee [(z_1 \vee z_2) \wedge (\bar{z}_1 \vee \bar{z}_2)]$$

which is equivalent to

$$\begin{aligned} & (x_1 \vee x_2 \vee \bar{y}_1 \vee y_2 \vee z_1 \vee z_2) \wedge (x_1 \vee x_2 \vee \bar{y}_1 \vee y_2 \vee \bar{z}_1 \vee \bar{z}_2) \\ & \wedge (x_1 \vee x_2 \vee y_1 \vee \bar{y}_2 \vee z_1 \vee z_2) \wedge (x_1 \vee x_2 \vee y_1 \vee \bar{y}_2 \vee \bar{z}_1 \vee \bar{z}_2) \\ & \wedge (\bar{x}_1 \vee \bar{x}_2 \vee \bar{y}_1 \vee y_2 \vee z_1 \vee z_2) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee \bar{y}_1 \vee y_2 \vee \bar{z}_1 \vee \bar{z}_2) \\ & \wedge (\bar{x}_1 \vee \bar{x}_2 \vee y_1 \vee \bar{y}_2 \vee z_1 \vee z_2) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee y_1 \vee \bar{y}_2 \vee \bar{z}_1 \vee \bar{z}_2). \end{aligned}$$

This establishes our claim.  $\square$

**Proposition 2.** *The 2-linear separability problem is NP-complete.*

*Proof.* The proof goes by reduction from reversible 6-SAT. Let  $(\xi_1^i \vee \xi_2^i \vee \xi_3^i \vee \xi_4^i \vee \xi_5^i \vee \xi_6^i)$  ( $i = 1, \dots, m$ ) be the clauses of the reversible 6-SAT problem, where  $\xi_k^i \in U = \{u_1, \dots, u_n, \bar{u}_1, \dots, \bar{u}_n\}$  ( $i = 1, \dots, m, k = 1, \dots, 6$ ). We first note that a reversible 6-SAT problem can be formulated as follows. Let  $x_1, \dots, x_n$  denote variables such that for each  $j$ ,  $6 < |x_j| < 7$ , with the interpretation that  $u_j$  is true if and only if  $x_j > 0$ . The clauses are then formulated as linear inequalities as shown in the following example. Consider the clause  $u_1 \vee \bar{u}_2 \vee u_3 \vee u_4 \vee \bar{u}_5 \vee u_6$ . This clause can be represented by the inequality  $x_1 - x_2 + x_3 + x_4 - x_5 + x_6 > -30$ . If the clause is false then  $x_1, x_3, x_4, x_6 < 0$  and  $x_2, x_5 > 0$ . Since  $6 < |x_j| < 7$ , we have  $x_1 - x_2 + x_3 + x_4 - x_5 + x_6 < -36$ . On the other hand, if the clause is true then  $x_1 - x_2 + x_3 + x_4 - x_5 + x_6 > -29$ . We now relate the constraint  $6 < |x_j| < 7$  to the 2-linear separability problem. Let  $e_j$  denote the unit  $n$ -vector with 1 in the  $j$ th position and 0's in all other positions. Together with every  $u_j$  we associate two points in the set  $P$ , namely,  $(\frac{1}{7})e_j$  and  $-(\frac{1}{7})e_j$ , and two points in the set  $Q$ , namely,  $(\frac{1}{6})e_j$  and  $-(\frac{1}{6})e_j$ . Also, we include the origin in the set  $P$ . Now, if there exist hyperplanes  $H_1$  and  $H_2$  as required (see the definitions of 2-linear separability), then  $x_0$  and  $y_0$  are positive since  $0 \in P$ . Without loss of generality, we may assume  $x_0 = y_0 = 1$ . Since  $(\frac{1}{7})e_j \in P$ , we have  $x_j < 7$  and  $y_j < 7$ . Similarly,  $-(\frac{1}{7})e_j \in P$  implies  $x_j > -7$  and  $y_j > -7$ . On the other hand, since  $(\frac{1}{6})e_j \in Q$ , we have either  $x_j > 6$  or  $y_j > 6$  and, since  $-(\frac{1}{6})e_j \in Q$ , either  $x_j < -6$  or  $y_j < -6$ . It follows that either  $-7 < x_j < -6$  and  $6 < y_j < 7$  or  $-7 < y_j < -6$  and  $6 < x_j < 7$ . This means that the  $y_j$ 's actually represent the negations of the corresponding  $x_j$ 's. Of course, we also include in the set  $P$  one point per each clause as in the following example. Given the clause  $u_1 \vee \bar{u}_2 \vee u_3 \vee u_4 \vee \bar{u}_5 \vee u_6$ , consider the inequality  $x_1 - x_2 + x_3 + x_4 - x_5 + x_6 > -30$ . This inequality can be stated as the requirement that the point

$(\frac{1}{30})(-1, 1, -1, -1, 1, -1, 0, \dots, 0)^T$  belongs to  $P$ . Notice that if this point belongs to  $P$  then we must also have  $y_1 - y_2 + y_3 + y_4 - y_5 + y_6 > -30$ . We have argued before that the  $y_j$ 's represent the negations of the corresponding  $x_j$ 's. However, since the formula is reversible, we may assume without loss of generality that the clause  $\bar{u}_1 \vee u_2 \vee \bar{u}_3 \vee \bar{u}_4 \vee u_5 \vee \bar{u}_6$  is also one of the conjuncts so the constraint  $y_1 - y_2 + y_3 + y_4 - y_5 + y_6 > -30$  does not affect the set of feasible solutions. To summarize the reduction, we associate with each variable two points in  $P$  and two points in  $Q$ . Also,  $0 \in P$ , and also with each clause we associate one more point in  $P$ . □

### 3. Separability with Many Lines in the Plane

We now turn to the  $k$ -polyhedral separability problem with a general  $k$ . The problem can be formulated as follows.

**Problem** ( $k$ -Polyhedral Separability). Given two sets of points with integer coordinates  $P = \{\pi^1, \dots, \pi^p\} \subset \mathbb{R}^d$  and  $Q = \{\rho^1, \dots, \rho^q\} \subset \mathbb{R}^d$ , and an integer  $k$ , recognize whether there exist  $k$  hyperplanes  $H_j = \{z: (x^j)^T z = x_0^j\}$  ( $x^j \in \mathbb{R}^d, x_0^j \in \mathbb{R}, j = 1, \dots, k$ ) that separate the sets  $P$  and  $Q$  through a boolean formula as follows. Associate with each hyperplane  $H_j$  a boolean variable  $\xi_j$ . The variable  $\xi_j$  is true at a point  $z$  if  $(x^j)^T z > x_0^j$  and false if  $(x^j)^T z < x_0^j$ . It is not defined at points lying on the hyperplane itself. A boolean formula  $\varphi = \varphi(\xi_1, \dots, \xi_k)$  separates the sets  $P$  and  $Q$  if  $\varphi$  is true at each of the points  $\pi^1, \dots, \pi^p$  and false at each of the points  $\rho^1, \dots, \rho^q$ .

Notice that we can distinguish here two types of problems, depending on whether the boolean formula is given or not. In the latter case we ask whether there exist a formula and hyperplanes that establish the separation. The NP-completeness of the case of a given formula (in general dimension with  $k$  fixed) follows from NP-completeness of the 2-linear separability problem. The other case can be proven NP-complete using similar methods and a characterization given below the 2-linear separability problem and is hence NP-complete in general dimension. It is interesting to consider the complexity of this problem also in fixed dimension. We first prove that already in two dimensions the problem with general  $k$  is NP-complete.

**Proposition 3.** *The hyperplanes  $H_1, \dots, H_k$  separate the sets  $P$  and  $Q$  in the sense of  $k$ -polyhedral separability through some boolean formula if and only if for every pair of points,  $\pi \in P$  and  $\rho \in Q$ , there exists an  $l$  ( $1 \leq l \leq k$ ) such that  $\pi$  and  $\rho$  lie on different sides of the hyperplane  $H_l$ .*

*Proof.* Given the hyperplanes  $H_1, \dots, H_k$ , consider the boolean variables  $\xi_1, \dots, \xi_k$  used in the definition of  $k$ -polyhedral separability. If two points  $z^1, z^2$  lie on the same side of each of the hyperplanes then the truth values of these boolean variables are the same for both of the points. If the hyperplanes separate

the sets  $P$  and  $Q$  in the sense of the definition then for any two points  $\pi^i, \rho^j$  there is at least one of the variables that has different truth values at  $\pi^i$  and  $\rho^j$ . This implies that at least one of the hyperplanes separates the two points. We now prove the converse. Suppose every two points  $\pi^i, \rho^j$  are separated by at least one hyperplane. Let  $\zeta$  be any point in  $P \cup Q$ . Let  $i_1, \dots, i_s$  denote the indices of the variables  $\xi_i$  which are true at  $\zeta$ , and let  $i_{s+1}, \dots, i_k$  denote the indices of those which are false at  $\zeta$  ( $0 \leq s \leq k$ ). Let

$$\varphi_\zeta = \varphi_\zeta(\xi_1, \dots, \xi_k) = \xi_{i_1} \wedge \dots \wedge \xi_{i_s} \wedge \bar{\xi}_{i_{s+1}} \wedge \dots \wedge \bar{\xi}_{i_k}$$

denote a boolean formula associated with  $\zeta$ . Obviously,  $\varphi_\zeta$  is true at  $\zeta$ . Consider the formula

$$\varphi = \bigvee_{\pi \in P} \varphi_\pi.$$

Obviously,  $\varphi$  is true at each point of  $P$ . On the other hand, for every  $\rho \in Q$ ,  $\varphi_\pi$  is false at  $\rho$  for every  $\pi \in P$  since at least one of the variables has different truth values at  $\pi$  and  $\rho$ . In other words, the formula

$$\neg\varphi = \bigwedge_{\pi \in P} \neg\varphi_\pi$$

is true at every  $\rho \in Q$ . This proves that the sets  $P$  and  $Q$  are separated in the sense of the definition through the formula  $\varphi$ . □

We now discuss the complexity of separability in the plane. We are interested in recognizing whether two sets of points in the plane can be separated by  $k$  straight lines. A related problem was considered in [7]:

**Problem (Point Covering).** Given a finite set of points in the plane and a number  $k$ , recognize whether there exist  $k$  straight lines such that each point lies on at least one of the lines.

The point covering problem was shown in [7] to be NP-complete. The proof can be adapted to establish the following:

**Proposition 4.** *The problem of recognizing whether two sets of points in the plane are separable by  $k$  lines is NP-complete.*

*Proof.* Membership in NP follows from the fact that if two sets are separable then there exist separating lines where the length of the binary representation of the coefficients is bounded by a polynomial in the length of the binary representation of the input points. The proof of completeness for NP will be established by the construction described in the remainder of the present section. □

Proposition 3 hints that the point covering problem is closely related to the separation problem with  $k$  lines. The construction here is in a sense an adaptation

of the construction of [7]. The reduction is from the 3-satisfiability problem. Consider a formula  $\varphi = E_1 \wedge \cdots \wedge E_m$  where  $E_j = x_j \vee y_j \vee z_j$ ,  $\{x_j, y_j, z_j\} \subset \{u_1, \bar{u}_1, \dots, u_n, \bar{u}_n\}$ ,  $j = 1, \dots, m$ . Without loss of generality, assume that none of the clauses contains both  $u_i$  and  $\bar{u}_i$  for any  $i$ . We shall construct a family of points and lines in the plane. Throughout the construction process, whenever a new point has to be picked (rather than be determined by the previously constructed objects), it is chosen to be in general position relative to the previously constructed objects. The exact sense of general position is explained later. We first describe the construction for the point covering problem. It is convenient to separate the process into steps:

1. The first step is to pick  $2(m+n)$  "locations" as follows. Together with every clause  $E_j$  ( $j = 1, \dots, m$ ), we associate a point  $\pi_j \in R^2$ . We also associate with each clause  $E_j$  another point  $\tau_j \in R^2$  whose role is explained later. Similarly, for every variable  $u_i$  ( $i = 1, \dots, n$ ), we associate a point  $\sigma_i \in R^2$  with  $u_i$  and a point  $\bar{\sigma}_i \in R^2$  with  $\bar{u}_i$ .
2. The second step is to determine lines  $L_{ij}$  and  $\bar{L}_{ij}$  as follows. For every  $i$  ( $i = 1, \dots, n$ ) and  $j$  ( $j = 1, \dots, m$ ), if  $u_i \in \{x_j, y_j, z_j\}$  then  $L_{ij}$  is the line determined by  $\sigma_i$  and  $\pi_j$ ; otherwise,  $L_{ij}$  is the line determined by  $\sigma_i$  and  $\tau_j$ . Analogously, if  $\bar{u}_i \in \{x_j, y_j, z_j\}$  then  $\bar{L}_{ij}$  is the line determined by  $\bar{\sigma}_i$  and  $\pi_j$ ; otherwise  $\bar{L}_{ij}$  is the line determined by  $\bar{\sigma}_i$  and  $\tau_j$ .
3. The third step is to determine for every  $i$  ( $i = 1, \dots, n$ ) a grid of  $m^2$  points as follows. For every  $j$  ( $j = 1, \dots, m$ ) and  $k$  ( $k = 1, \dots, m$ ), denote by  $\pi_{jk}^i$  the point of intersection of the line  $L_{ij}$  with the line  $\bar{L}_{ik}$ .

We require that all the points of the type  $\pi_{jk}^i$  and  $\pi_j$  be in general position (subject to the rules above) in the sense that, except for the lines of the types  $L_{ij}$  and  $\bar{L}_{ij}$ , no other line in the plane contains more than two of these points, and none of these points lies on any line of those types unless it is required to by definition. It is easy to satisfy these requirements, for example, by small perturbations. Note that the locations are constructed so that for every  $j$  ( $j = 1, \dots, m$ ), the location  $\pi_j$  lies on a line  $L_{ik}$  if and only if  $j = k$  and  $u_i \in \{x_j, y_j, z_j\}$ ; the location  $\pi_j$  lies on a line  $\bar{L}_{ik}$  if and only if  $j = k$  and  $\bar{u}_i \in \{x_j, y_j, z_j\}$ . It turns out that the minimum number of lines required to cover the locations  $\pi_{jk}^i$  (in the sense that each location belongs to at least one line) is precisely  $mn$ . The covering lines have to be of the types  $L_{ij}$  and  $\bar{L}_{ij}$ , where for each  $i$  a unique type of line has to be chosen, that is, either  $\{L_{i1}, \dots, L_{im}\}$  or  $\{\bar{L}_{i1}, \dots, \bar{L}_{im}\}$ . It follows that the entire collection of locations (that is, including the locations of the type  $\pi_j$ ) can be covered by  $mn$  lines if and only if the given formula is satisfiable. A satisfying assignment corresponds to the choice of type of lines for each of the "grids." It is shown in [7] that the locations can be constructed so that their coordinates are bounded by a polynomial in  $m$  and  $n$ . So far we have essentially repeated the reduction of 3-satisfiability to point covering.

For the separability problem we do as follows. We have to construct two sets of points  $P$  and  $Q$  and consider the problem of separating them with a given number of lines. We use the convention that points denoted with the letter  $\pi$  belong to the set  $P$  while those denoted with the letter  $\rho$  belong to the set  $Q$ .

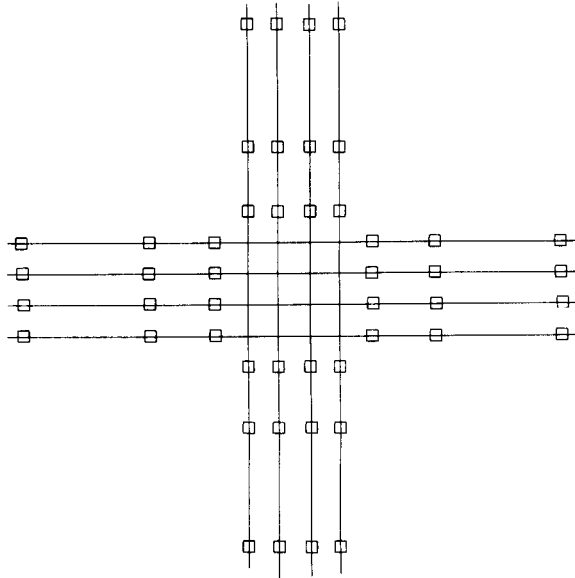
Intuitively, the covering problem can be related to the separability problem by splitting points of the covering problem into pairs of points to be separated. However, the requirement in the separation problem is that every point of  $P$  be separated from every point of  $Q$  and this necessitates several modifications and additions to the construction.

Suppose we have constructed  $N' = m + nm^2$  points  $\pi_1, \dots, \pi_N$  as explained above for the reduction of 3-satisfiability to the covering problem. We split each point  $\pi_j$  into two points  $\pi_j$  and  $\rho_j$  close to the previous  $\pi_j$ . A solution to the covering problem will yield a set of lines that separate every pair  $(\pi_j, \rho_j)$  but not necessarily every pair  $(\pi_j, \rho_k)$ . To overcome this difficulty, we use auxiliary sets of points  $P^0$  and  $Q^0$  such that a family of lines separating  $P^0$  and  $Q^0$  also separates every  $\pi_j$  from every  $\rho_k$  such that  $k \neq j$ .

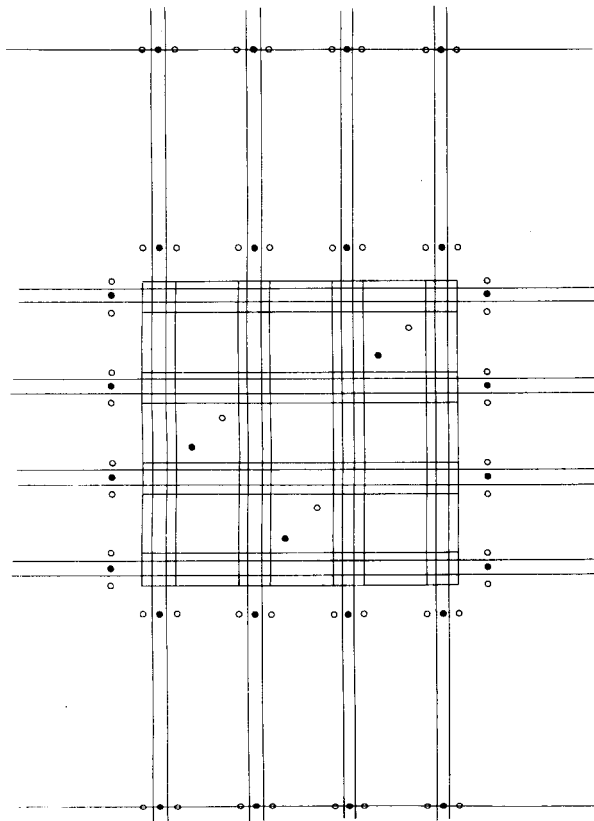
Without loss of generality, suppose the coordinates of all the points are divisible by 6 and are pairwise distinct. We now construct auxiliary points as follows. Let  $N = 2N'$ . Let  $U = 6u$  denote an upper bound on the absolute value of the coordinates of any point  $\pi_j$ . Also, define a sequence  $\{d_j\}$  by setting  $d_1 = U + 6$  and  $d_{j+1} = 3d_j + 2U + 5$ . The set  $P^0$  consists of all points of the forms  $(\pm(6k-3), \pm d_j)$  or  $(\pm d_j, \pm(6k-3))$ , where  $1 \leq k \leq u+1$  and  $1 \leq j \leq N$ . We also construct the set  $Q^0$  of all the points of one of the following forms:  $(\pm(6k-4), \pm d_j)$ ,  $(\pm(6k-2), \pm d_j)$ ,  $(\pm d_j, \pm(6k-4))$ , and  $(\pm d_j, \pm(6k-2))$  with  $k$  and  $j$  as above (see Figs. 1 and 2).

We refer below to groups of intervals which we call the *auxiliary intervals*. A group of auxiliary intervals is of one of the four types listed below. Each group contains  $4N$  intervals and includes for a fixed  $k$  ( $k = 1, \dots, u+1$ ) all the intervals of one of the forms defining the type. The types are as follows:

$$\begin{aligned}
 T_{h+}(k) &= \{((6k-4), d_j), (6k-3), d_j): j = 1, \dots, N\} \\
 &\cup \{((6k-3), d_j), (6k-2), d_j): j = 1, \dots, N\} \\
 &\cup \{((6k-4), -d_j), (6k-3), -d_j): j = 1, \dots, N\} \\
 &\cup \{((6k-3), -d_j), (6k-2), -d_j): j = 1, \dots, N\} \quad (k = 1, \dots, u+1), \\
 T_{h-}(k) &= \{((-6k+4), d_j), (-6k+3), d_j): j = 1, \dots, N\} \\
 &\cup \{((-6k+3), d_j), (-6k+2), d_j): j = 1, \dots, N\} \\
 &\cup \{((-6k+4), -d_j), (-6k+3), -d_j): j = 1, \dots, N\} \\
 &\cup \{((-6k+3), -d_j), (-6k+2), -d_j): j = 1, \dots, N\} \quad (k = 1, \dots, u+1), \\
 T_{v+}(k) &= \{((d_j, 6k-4), (d_j, 6k-3)): j = 1, \dots, N\} \\
 &\cup \{((d_j, 6k-3), (d_j, 6k-2)): j = 1, \dots, N\} \\
 &\cup \{((-d_j, 6k-4), (-d_j, 6k-3)): j = 1, \dots, N\} \\
 &\cup \{((-d_j, 6k-3), (-d_j, 6k-2)): j = 1, \dots, N\} \quad (k = 1, \dots, u+1), \\
 T_{v-}(k) &= \{((d_j, -6k-4), (d_j, -6k+3)): j = 1, \dots, N\} \\
 &\cup \{((d_j, -6k-3), (d_j, -6k+2)): j = 1, \dots, N\} \\
 &\cup \{((-d_j, -6k+4), (-d_j, -6k+3)): j = 1, \dots, N\} \\
 &\cup \{((-d_j, -6k+3), (-d_j, -6k+2)): j = 1, \dots, N\} \quad (k = 1, \dots, u+1).
 \end{aligned}$$



**Fig. 1**



**Fig. 2**



Note that altogether we have  $4(u + 1)$  pairwise disjoint groups, each consisting of  $4N$  auxiliary intervals.

**Proposition 5.**

- (i)  $8(u + 1)$  straight lines are both necessary and sufficient for separating the sets  $P^0$  and  $Q^0$ .
- (ii) For every family of  $8(u + 1)$  lines separating  $P^0$  from  $Q^0$  and for each of the  $4(u + 1)$  groups of auxiliary intervals  $T_{h+}(k)$ ,  $t_{h-}(k)$ ,  $T_{v+}(k)$ , and  $T_{v-}(k)$  ( $k = 1, \dots, u + 1$ ), all the  $4N$  auxiliary intervals in the group must be intersected by the same two lines of the family.

*Proof.* (i) First, the  $8(u + 1)$  lines  $\{(x, y): x = \pm(6k - 3.5)\}$ ,  $\{(x, y): x = \pm(6k - 2.5)\}$ ,  $\{(x, y): y = \pm(6k - 3.5)\}$ , and  $\{(x, y): y = \pm(6k - 2.5)\}$  ( $k = 1, \dots, u + 1$ ) separate these points so  $8(u + 1)$  lines suffice. On the other hand, the union of these groups of intervals consists of  $16N(u + 1)$  intervals all of which must be intersected. It is easy to see that there exists no straight line which intersects more than  $2N$  of these intervals and hence at least  $8(u + 1)$  lines are required.

(ii) Altogether we have  $16N(u + 1)$  intervals which must be intersected by  $8(u + 1)$  lines. No line can intersect more than  $2N$  intervals. Thus, each line must intersect exactly  $2N$  auxiliary intervals. It is easy to verify that if a straight line intersects  $2N$  auxiliary intervals then all these intersected intervals must either be parallel to the  $x$ -axis and have pairwise distinct  $y$ -coordinates or be parallel to the  $y$ -axis and have pairwise distinct  $x$ -coordinates. It is also easy to check that if a straight line intersects  $2N$  intervals then the same value of the index  $k$  must be involved in the specification of all of these intervals as described above. This completes the proof. □

The roles of the sets  $P^0$  and  $Q^0$  will become clear later. Recall that we have constructed points  $\pi_j = (\xi_j, \eta_j)$  for the covering problem and have assumed that the coordinates of these points are divisible by 6 and are pairwise distinct. We now split each such point in two. Before we continue with the construction it is essential first to prove the following fact about the point covering problem.

**Proposition 6.** *Let  $S$  be a set of points in the plane whose coordinates are integers between  $-K$  and  $K$ . Suppose the minimum number of lines required for covering all the points of  $S$  is  $s$ . For each point  $p \in S$ , let  $I_p$  denote a small line segment of length  $\epsilon < 1/(12K + 6)$  centered at  $p$ . Under these conditions, the minimum number of lines required to intersect all the segments  $I_p$  is also equal to  $s$ .*

*Proof.* It is trivial that  $s$  lines suffice. Moreover, if all the line segments are sufficiently small then  $s$  lines are necessary. It remains to show that  $\epsilon < 1/(12K + 6)$  is sufficiently small. It suffices to show that for every three noncollinear points  $p_1, p_2, p_3 \in S$ , every three points  $p'_i \in I_{p_i}$  ( $i = 1, 2, 3$ ) are noncollinear.

Suppose  $p_i = (a_i, b_i)$  ( $i = 1, 2, 3$ ) are not collinear. Since the coordinates are integer, we may assume without loss of generality that

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \geq 1.$$

Consider points of the form  $p'_i = (a_i + \varepsilon_{1i}, b_i + \varepsilon_{2i})$  ( $i = 1, 2, 3$ ). In order for these points to be noncollinear, it is necessary and sufficient that

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 + \varepsilon_{11} & a_2 + \varepsilon_{12} & a_3 + \varepsilon_{13} \\ b_1 + \varepsilon_{21} & b_2 + \varepsilon_{22} & b_3 + \varepsilon_{23} \end{vmatrix} \neq 0.$$

If  $|\varepsilon_{ij}| \leq \varepsilon$  then this determinant is not less than  $1 - 6\varepsilon^2 - 12K\varepsilon$  and hence positive. This completes the proof. □

Let us now fix

$$\varepsilon \leq \frac{1}{12U + 6}.$$

Without loss of generality, assume none of the lines participating in the solution of the point covering problem has slope 1. For each  $j$ , let  $\rho_j = (\xi_j - \varepsilon, \eta_j - \varepsilon)$  and let us redefine  $\pi_j = (\xi_j + \varepsilon, \eta_j + \varepsilon)$ . Denote the set of these new points  $\pi_j$  by  $P^1$  and the set of the points  $\rho_j$  by  $Q^1$ .

**Proposition 7.** *For any set of  $8(u + 1)$  lines separating the set  $P^0$  and  $Q^0$ , and for every  $j \neq k$  ( $1 \leq j, k \leq N$ ), the point  $\pi_j$  is separated from all the points  $\rho_k$  by at least one of the lines in the family.*

*Proof.* The proof follows from the construction. □

Consider the problem of separating the sets  $P = P^0 \cup P^1$  and  $Q = Q^0 \cup Q^1$  with a minimum number of lines. We know that it takes  $8(u + 1)$  lines to separate  $P^0$  from  $Q^0$ . Moreover, such lines also separate  $P^0$  from  $Q^1$  and  $Q^0$  from  $P^1$ . Also, they separate points of  $P^1$  from points of  $Q^1$  provided they have distinct indices. Thus, it remains to separate points of  $P^1$  from points of  $Q^1$  with the same index. The problem of separating such pairs is equivalent to the covering problem. We know it takes at least  $mn$  lines in any case. Moreover,  $mn$  lines suffice if and only if the given formula is satisfiable. It is important to notice that the  $8(u + 1)$  lines that separate  $P^0$  and  $Q^0$  are of no help in separating points of  $P^1$  and  $Q^1$  with the same index. To see this, notice that it follows from our choice of the sequence  $\{d_j\}$  that if a straight line intersects an interval of the form  $(\pi_j, \rho_j)$  then it cannot intersect more than two of the auxiliary intervals. The separation problem is related to the satisfiability problem as follows. We will prove that the given formula is satisfiable if and only if  $8(u + 1) + mn$  lines suffice for separating the sets  $P$  and  $Q$ .

**Proposition 8.**

- (i) *If the formula  $\varphi$  is satisfiable then the set  $P$  can be separated from the set  $Q$  with  $8(u + 1) + mn$  lines.*
- (ii) *If  $L \leq 8(u + 1) + mn$  lines separate  $P$  from  $Q$  then necessarily  $8(u + 1)$  of them separate  $P^0$  from  $Q^0$  and  $mn$  of them separate  $P^1$  from  $Q^1$ , so  $L = 8(u + 1) + mn$  and the formula  $\varphi$  is satisfiable.*

*Proof.* The proof of (i) is obvious. We now prove (ii). Suppose we designate  $l$  of the  $L$  lines to intersect intervals of the type  $(\pi_j, \rho_j)$ . Since there are only  $N'$  such intervals, we may assume without loss of generality that  $l \leq N'$ . Each such line can intersect at most two of the auxiliary intervals. Thus, the designated lines intersect at most  $2l$  auxiliary intervals. We are therefore left with at least  $16N(u + 1) - 2l$  auxiliary intervals which require at least

$$\left\lceil \frac{16N(u + 1) - 2l}{2N} \right\rceil = 8(u + 1)$$

lines to intersect all of them. But the auxiliary intervals require this number of lines to intersect all of them in any case. Moreover, no line can intersect more than  $2N$  auxiliary intervals and  $2l \leq N$ . It follows that each of these  $8(u + 1)$  lines must intersect at least  $N$  auxiliary intervals and hence cannot intersect any interval of the form  $(\pi_j, \rho_j)$ . However, we know that the latter intervals require at least  $mn$  lines to intersect all of them and this is feasible if and only if the formula  $\varphi$  is satisfiable. This completes the proof. □

The proof of Proposition 4 is now established.

**4. Separability with a Fixed Number of Hyperplanes in a Fixed Dimension**

It is interesting to note that if both the dimension of the space and the number  $k$  are fixed then the  $k$ -polyhedral separability problem is solvable in polynomial time. This is what we prove in the present section.

Recall that separation throughout this paper is in the strong sense, namely, two sets of points  $A, B \subset R^d$  are separated by the hyperplane  $H = \{z \in R^d : z^T x = x_0\}$  if for every  $a \in A, a^T x < x_0$  and for every  $b \in B, b^T x > x_0$ , or vice versa.

**Proposition 9.** *Suppose  $A$  and  $B$  are sets of points in  $R^d$  with integer coordinates, and suppose there exists a hyperplane  $H = \{z \in R^d : y^T z = y_0\}$  that separates  $A$  from  $B$ , assuming  $y^T a < y_0$  for  $a \in A$ . Under these conditions, there exist a hyperplane  $H = \{z \in R^d : z^T x = x_0\}$ , a positive rational number  $r$ , and integers  $j_A, j_B$  ( $j_A, j_B \geq 1, j_A + j_B \leq d + 1$ ) such that:*

- (i) *For every  $a \in A, x^T a \leq x_0 - r$ .*

- (ii) For every  $b \in B$ ,  $x^T b \geq x_0 + r$ .  
 (iii) For at least  $j_A$  points  $a \in A$ ,  $x^T a = x_0 - r$ , and for at least  $j_B$  points  $b \in B$ ,  $x^T b = x_0 + r$ .

*Proof.* Consider the following linear programming problem where the unknowns are the vector  $x \in R^d$  and the scalars  $x_0$  and  $r$ :

$$\begin{aligned}
 (P) \quad & \text{Maximize} \quad r \\
 & \text{subject to} \quad x^T a \leq x_0 - r \quad (a \in A), \\
 & \quad \quad \quad x^T b \geq x_0 + r \quad (b \in B), \\
 & \quad \quad \quad -1 \leq x_j \leq 1 \quad (j = 1, \dots, d), \\
 & \quad \quad \quad r \geq 0.
 \end{aligned}$$

First, note that the existence of a feasible solution for  $(P)$  with  $r > 0$  follows from the assumption of existence of a separating hyperplane. Also,  $(P)$  is obviously bounded and hence has an optimal solution. Moreover, there exists a basic optimal solution, that is, an optimal solution where  $d + 2$  linearly independent constraints are satisfied as equalities. At any optimal solution at least one equality of the form  $x_j = \pm 1$  ( $1 \leq j \leq d$ ) has to hold, since otherwise we could increase  $r$  by multiplying all the inequalities by some number greater than 1. Similarly, at least one of the constraints corresponding to  $A$  and at least one of the constraints corresponding to  $B$  have to be satisfied as equalities. This completes the proof.  $\square$

**Proposition 10.** For every fixed  $k$  and  $d$ , the  $k$ -polyhedral separability problem in  $R^d$  can be solved in polynomial time.

*Proof.* If two sets  $P, Q \subset R^d$  are separable with  $k$  hyperplanes then there exist  $k$  pairs of complementary subsets  $A_i, B_i \subset P \cup Q$  (that is,  $A_i \cup B_i = P \cup Q$ ,  $i = 1, \dots, k$ ) and  $k$  hyperplanes  $H_i$  ( $i = 1, \dots, k$ ) such that  $H_i$  separates  $A_i$  from  $B_i$ . From the proof of Proposition 9, it follows that there exist such hyperplanes that satisfy equalities as stated in that proposition, and, moreover, the separating hyperplanes can be chosen from a finite set. Each of the candidate hyperplanes is determined (see the proof of Proposition 9) by some set of at most  $d + 1$  points, together with a choice of at most  $d$  equalities  $x_j = \pm 1$ . The number of such sets of at most  $d + 1$  points is of course polynomial in the cardinality of  $P \cup Q$ . Thus, the number of combinations of  $k$  such sets is also polynomial. It follows that we can enumerate in polynomial time all the relevant configurations of hyperplanes. Furthermore, it takes polynomial time to check whether a given configuration actually separates  $P$  from  $Q$  (see Proposition 3). This establishes the proof.  $\square$

Note that Proposition 10 does not rely on the fact that the linear programming problem is solvable in polynomial time. In fact, here we have polynomial time in the strong sense that it is valid also under the real number model of computation.

## Acknowledgment

The author is grateful to David Haussler for suggesting the general problem discussed in this paper, and to Arie Tamir for several helpful comments.

## References

1. A. Aggarwal, H. Booth, J. O'Rourke, S. Suri, and C. K. Yap, Finding minimal convex nested polygons, *Proceedings of the 1st ACM Symposium on Computational Geometry* (1985), ACM, New York, 1985, pp. 296–304.
2. A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth, Classifying learnable geometric concepts with the Vapnik–Chervonenkis dimension, *Proceedings of the 18th Annual ACM Symposium on Theory of Computing* (1986), ACM, New York, 1986, pp. 273–282.
3. H. Edelsbrunner and F. P. Preparata, Minimum polygonal separation, Report No. UIUCDCS-R-86-1272, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, May 1986.
4. C. E. Kim and T. A. Anderson, Digital disks and a digital compactness measure, *Proceedings of the 16th Annual ACM Symposium on Theory of Computing* (1984), ACM, New York, 1984, pp. 117–124.
5. N. Megiddo, Linear-time algorithms for linear programming in  $R^3$  and related problems, *SIAM Journal on Computing* **12** (1983), 759–776.
6. N. Megiddo, Linear programming in linear time when the dimension is fixed, *Journal of the Association for Computing Machinery* **31** (1984), 114–127.
7. N. Megiddo and A. Tamir, On the complexity of locating linear facilities in the plane, *Operations Research Letters* **1** (1982), 194–197.
8. J. O'Rourke, S. R. Kosaraju, and N. Megiddo, Computing Circular Separability, *Discrete Comput. Geom.* **1** (1986), 105–113.

Received August 12, 1986, and in revised form June 26, 1987.