1

of/de

T

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NRS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

PIONEERS IN METHYLENE BLUE TESTING SINCE 1974

1544 (NAT) (REAL) 5 BURGESSLE MM 05317 USA
I 01 435 100 FAX 012 435 7877 TEL 5100000000

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Separating Convex Sets

by

Ralph Patrick Boland

**A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Master of Computer Science**

**School of Computer Science
Carleton University, Ottawa, Ontario**

October 15, 1991

© copyright 1991, Ralph Patrick Boland

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-75969-0

Canada

The undersigned recommend to the Faculty of Graduate Studies and Research
acceptance of the thesis

"Separating Convex Sets"

submitted by Ralph Patrick Boland, B.Sc.
in partial fulfilment of the requirements for
the degree of Master of Computer Science

Thesis Supervisor

Director, School of Computer Science

Carleton University
December 6, 1991

ABSTRACT

Given two disjoint convex sets A and B in E^d , a hyperplane h in E^d separates them if A lies on one of the half spaces defined by h while B lies on the complementary half space. Given a collection F of convex sets in E^d we say F is separated by a set of hyperplanes H if every pair of elements of F is separated by some hyperplane of H . In this thesis we deal mainly with separability of point sets. Let $f(n,d)$ be the minimum number of hyperplanes always sufficient and occasionally necessary to separate n points in general position in E^d . We prove that $\lceil \frac{n-1}{d} \rceil \leq f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ and that $f(n,2) = \lceil \frac{n}{2} \rceil$. We present an algorithm to generate a separating family of hyperplanes in E^d with expected time complexity of $O(n \log(n))$ and worst case time complexity of $O(n^2)$. In dimension 2 the worst case time complexity can be reduced to $O(n \log(n))$.

We also show that out of a collection of n circles there is always one circle that is separable from $\lfloor \frac{n}{3} \rfloor - 2$.

ACKNOWLEDGEMENTS

I would like to thank Jorge Urrutia, my thesis supervisor, not only for his considerable efforts in that role but also for all of his assistance during my stay in Mexico City. I am also very grateful to his wife, Margaret, for proof reading this thesis.

I would also like to thank my friends in Ottawa and in Mexico City for their encouragement and support.

Finally, I would like to thank my parents for always being there when they were needed.

TABLE OF CONTENTS

1	Introduction	1
2	Separating Convex Sets on the Plane	5
2.1	Separation of a Collection of Convex Sets by a Set of Lines.....	8
2.2	Separation of One Convex Set from Many Convex Sets.	10
3	Separating Circles and Line Segments.....	14
3.1	Upper Bounds on Separating Circles	15
3.2	Separating one Circle from Many Circles.....	17
3.3	Upper Bounds on Separating Line Segments.	29
3.4	Separating one Line Segment from Many.	32
4	Separating Points	36
4.1	Determination of Lower Bounds for $f(n,d)$	38
4.2	A Simple Point Separation Algorithm	40
4.3	Separating Points by Lines on the Plane	44
4.4	Separating Points by Hyperplanes in E^d	47
4.4.1	Extending Algorithm 2-Separate to Higher Dimensions	48
4.4.2	Efficient k -Partitioning of n Points in E^d where $k \geq d$	52

4.5	Separating points in efficiently in E^2 and E^3	62
4.6	Data Structures for Retrieval of Separating Hyperplanes and Lines.....	63
4.6.1	A Data Structure for Retrieval of Separating Hyperplanes.....	63
4.6.2	A Data Structure for Retrieval of Separating Lines	65
5	Summary and Open Problems	70
5.1	Some Open Problems	74

1 Introduction

In this thesis we present some new results concerning the separability of convex sets in Euclidean vector spaces. Let E^d represent d -dimensional Euclidean vector space. Given two disjoint convex sets A and B in E^d , a hyperplane h in E^d separates them if A lies on one of the half spaces defined by h while B lies on the complementary half space. Given a collection F of convex sets in E^d , we say F is separated by a set of hyperplanes H if, for each possible pair of elements of F , there is a hyperplane of H that separates them. Given an element $A \in F$ and a subcollection S of F we say that a hyperplane h separates A from S if h separates A from each element of S .

Some of our new results concern the separability of point sets in E^d . These are the first significant results on separability dealing with dimensions higher than two. Let $f(n,d)$ be the minimum number of hyperplanes always sufficient and occasionally necessary to separate n points in E^d where $n \geq d$. Then in E^2 (i.e. the plane) we prove that $f(n,2) = \lceil \frac{n}{2} \rceil$. For dimensions higher than two we prove that $\lceil \frac{n-1}{d} \rceil \leq f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d odd and $\lceil \frac{n}{d} \rceil \leq f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d even. The upper bounds of $f(n,d)$ are determined by construction: we present algorithms that generate a solution within these upper bounds. The algorithms are efficient, ranging in complexity from $O(n \log(n))$ to $O(n^2)$. The $O(n^2)$ algorithm is very easy to implement and very fast in practice, having an expected complexity of $O(n \log(n))$.

We then develop a data structure for storing the hyperplanes that separate a set of points so that retrieving a hyperplane, in the data structure, that separates two given points can be found in worst case time $O(n)$ and expected case time $O(\log(n))$. We also show that in

E^2 we can construct an even more efficient data structure that allows us to retrieve a hyperplane that separates two given points in worst case time $O(\log(n))$ and expected case time $O(1)$.

We also discuss the possibility of reducing the upper bound on $f(n,d)$ in general and of $f(n,4)$ in particular. In the process we shed some light on a well known problem in E^4 .

In [CCUZ] it is shown that for any set of n circles there is always a line that separates one circle from at least $\lceil \frac{n}{4} \rceil - 1$. We improve this result to $\lfloor \frac{n}{3} \rfloor - 2$. The new result is achieved by modifications to the proof given in the original paper.

As well as presenting our new results we provide a survey on the results to date on separability of convex sets. These results deal only with the planar case. In the process we expand some of the details and prove some lemmas that were left for the reader to verify.

In Chapter two we survey the results on the separation of arbitrary convex sets in the plane. We first present a result by [Tver] which shows that for any collection of five convex sets that there is one that can be separated by a line from at least two.

We next present a result by [Feje] and later by [ERS] and [UrZa] which proves that in the plane any set of n disjoint convex sets can be separated by a set of at most $3n-6$ lines; a result that leads directly from the properties of planar graphs.

We then consider the problem of separating at least one element of a collection F of convex sets from as many elements of F as possible. (We show in Chapter three that it is not generally possible to find four elements of F such that two of them can be separated by the line from the other two.) We then show that for any collection of n disjoint

convex sets F on the plane there exists a line that separates one convex set of F from a subcollection of F containing at least $\lceil \frac{n}{12} \rceil$ elements.

In Chapter three the separability of circles in the plane and the separation of line segments in the plane are discussed. A collection of n disjoint circles C_n is presented such that no line exists that has two circles from C_n on each side of it. It is also shown that for any collection C_n of n disjoint circles that one can always find a line L that separates one circle of C_n from a subcollection of C_n containing at least $\lfloor \frac{n}{3} \rfloor - 2$ circles but that for some collections of n circles C_n there does not exist a line λ that separates one circle of C_n from any subcollection of C_n containing $\lceil \frac{n}{2} \rceil$ or more circles.

The results on line segments are then surveyed. A configuration of n disjoint line segments T_n is presented such that no line exists that has two segments from T_n on each side of it. Then a configuration of $3n$ line segments J_n is presented such that no line can separate one element of J_n from a subcollection of J_n containing more than $n+1$ line segments. It is also shown, however, that for any collection S of n disjoint line segments that one can always find a line that separates one element of S from a subcollection of S containing at least $\lceil \frac{n-1}{4} \rceil + 1$ segments. If the elements of S can be extended to be semilines so that they remain disjoint then it is shown that there is an element of S that can be separated by a line from a subcollection of S which contains at least $\lceil \frac{n}{3} \rceil + 1$ elements.

The results of this chapter remain valid even if the convex set are permitted to be adjacent but must still have pairwise disjoint relative interiors.

In Chapter four we present the results on the separability of point sets in E^d . We prove the results stated above on the separability of point sets. We then consider how these results may be improved. We show a connection between the separability of point sets in

E^d and the existence of certain hamiltonian paths in hypercube graphs of dimension d . We consider the interesting idea of "General Position Curves" and how it relates to separating points in E^d .

Finally, in Chapter 5, after summarizing the results we discuss the open problems in this field and some others encountered in Chapter four.

2 Separating Convex Sets on the Plane

In this chapter we survey the separation of convex sets in the plane. The convex sets considered here are always compact (i.e. closed and bounded) but are otherwise arbitrary. We first present a result by [Feje] that any collection F of n convex sets in the plane can be separated by a set of at most $3n-6$ lines. We then use this result to prove that there is a line that separates an element of F from a subcollection of F containing at least $\lceil \frac{n}{12} \rceil$ elements. This proof is from [CCUZ].

This result is an improvement on a result by [Tver] which proved an upper bound of $R(n) + n - 1$, where $R(n)$ is a Ramsey number. Tverberg determines this result by first proving the following result. (We have simplified the proof somewhat.)

Theorem. 2.1. Let $F(n)$ be the smallest integer such that for any collection of $F(n)$ disjoint compact convex sets one can be separated from n . Then $F(2) = 5$.

Proof. To show that $F(2) \geq 5$ we consider three congruent circles mutually tangent and surrounding a fourth smaller circle which is tangent to all three outer circles. It remains to show that $F(2) \leq 5$.

Assume that $n \geq 6$ (possibly infinite). Then there is a collection of disjoint compact convex sets $C = C_1, \dots, C_5$ such that every line separating two of the sets of C intersects the other three. We show that no such C exists.

Let λ' and λ'' be the unique pair of lines which are tangent to both C_1 and C_2 and also separates them. Then $\lambda \setminus \{C_1, C_2\}$ consists of three components X' , Y' , and Z' which we label such that Y' is between X' and Z' and such that X' is closer to C_1

than to C_2 . $\lambda' \setminus \{C_1, C_2\}$ similarly consists of three components X'' , Y'' , and Z'' labeled such that Y'' is between X'' and Z'' and such that X'' is closer to C_1 than to C_2 .

Now $C_3 \cap Y' = \emptyset$ since if C_3 intersects Y' , say between the part of λ' between $\lambda' \cap \lambda''$ and C_1 , then any line separating C_3 from C_1 avoids C_2 , which contradicts the definition of C . We assume, without loss of generality, that C_3 intersects X' (and therefore does not intersect Z'). Similarly C_3 intersects X'' or Z'' but not both. However if C_3 intersects X'' then any line separating C_3 from C_1 cannot intersect C_2 and therefore C_3 intersects X' and Z'' . Similarly C_4 intersects either X' and Z'' or X'' and Z' . The same is true for C_5 . Therefore, up to symmetry we have two cases:

- 1) C_3, C_4 , and C_5 intersect X' and Z'' .
- 2) C_3 and C_4 intersect X' and Z'' , while C_5 intersects X'' and Z' .

We need not consider Case 1 since in this case, if we instead choose λ' and λ'' as the lines tangent to both C_3 and C_4 and then relabel C_3 and C_4 by C_1 and C_2 and vice versa, we get Case 2.

We now need consider only Case 2. See Figure 2.1. We assume, without loss of generality, that C_3 is closer to $\lambda' \cap \lambda''$ than is C_4 . If we now reselect λ' and λ'' as the lines tangent to both C_3 and C_5 and then relabel C_3 and C_5 by C_1 and C_2 and vice versa, we get that C_3 intersects λ' along Y' . (We could also have chosen C_5 or Y'' or both). Then any line separating C_3 from C_1 avoids C_2 which contradicts the definition of C .

♦

Two convex sets can always be separated from one out of a collection of five convex sets.

Figure 2.1

Theorem 2.1 represents the largest case, that we know of, for which the value of $F(n)$ is known. It is apparent from the proof that determining $F(n)$ for $n > 2$ is a difficult problem.

2.1 Separation of a Collection of Convex Sets by a Set of Lines.

We now show that any collection F of n convex sets in the plane can be separated by a set of at most $3n-6$ lines. This result was first shown in [Feje] and later by [ERS] and [UrZa]. This result is a surprise since, as we shall see in Chapter three, separating convex sets is in general difficult and even finding a line that separates any two elements of F from any other two elements of F is in general not possible.

Theorem 2.2 For any collection F of n disjoint bounded and closed convex sets on the plane there exists a partitioning π of the plane using k line segments or semilines $R_1, \dots, R_k$, with $k \leq 3n-6$, such that every element in F lies on a different face of π and every element R_i of π lies on the boundary of exactly two of the faces of π that contain elements of F . See Figure 2.2.

Proof. Let $F = \{C_1, C_2, \dots, C_n\}$ be any collection of mutually disjoint bounded and closed convex sets in the plane all contained within a triangle T . Inflate all the sets C_i , $1 \leq i \leq n$, to get a collection of n compact and closed convex sets $F^* = \{D_1, \dots, D_n\}$ inside T , such that:

- a) $D_i \supseteq C_i$, $1 \leq i \leq n$
- b) Any intersecting pair of elements in F^* intersect only along some portion of their boundary.
- c) No D_i can be extended so as to increase the number of intersecting pairs of F^* .

See Figure 2.3.

Separating n convex sets with at most $3n-6$ lines

Figure 2.2.

inflation of convex sets until they are adjacent

Figure 2.3

For any $i, j, 1 \leq i < j \leq n$, such that D_i is tangent to D_j let $\lambda_{i,j}$ be the line separating D_i from D_j . Let L be the set of all such lines. Clearly L is a set of lines that separates F^* and thus also separates F . We now count the elements of L . Construct the following graph G . Let the vertices of G correspond to the elements of F^* . Two vertices of G are adjacent if their corresponding elements of F^* are tangent. Clearly $E(G) = |L|$. Also G is clearly planar and therefore $E(G) = |L| \leq 3n-6$.

♦

Corollary 2.3. For $n \geq 3$ there are collections of n convex sets F such that $3n-6$ lines are necessary to separate F .

Proof. Construct a collection $F = \{C_1, \dots, C_n\}$ as follows: Let C_1, C_2, C_3 be any three mutually tangent circles. Clearly three uniquely determined lines are needed to separate these circles. Let $C_i, 4 \leq i \leq n$ be the circle which is mutually tangent to the circles $\{C_{i-3}, C_{i-2}, C_{i-1}\}$ and which does not overlap these circles. Clearly three uniquely determined lines are needed to separate C_i from the three circles tangent to C_i and these lines cannot also separate any pair of circles from $\{C_1, \dots, C_{i-1}\}$. Therefore $3n-6$ lines are needed to separate F .

2.2 Separation of One Convex Set from Many Convex Sets.

We now determine a lower bound on the number of convex sets that can always be separated from one out of a collection of n convex sets. This result is from [CCUZ].

Theorem 2.4. For any collection F of n disjoint convex sets on the plane, there exists a line λ that separates an element $A \in F$ from a subcollection of F containing at least $\lceil \frac{n}{12} \rceil$ elements.

For this proof we need the following definitions and lemma.

It is well known that for any pair of disjoint bounded convex plane polygons P and Q , there exists a unique pair of lines λ_1 and λ_2 such that each of them is tangent to both P and Q and also separates them. λ_1 and λ_2 also exist if P and Q are unbounded unless there are two unbounded edges (semilines), one of P and one of Q , that are parallel and either of them determines a line that separates P from Q . We refer to λ_1 , λ_2 , if they exist, as the *crossing support lines* for P, Q . For any line segment or semiline e , let $L(e)$ denote the line containing e .

Lemma 2.5. Let P and Q be any two disjoint convex plane polygons, possibly unbounded. Then there exists at least one edge e of P or Q such that $L(e)$ separates P from Q .

Proof. We assume that P and Q determine a pair of crossing supporting lines, say λ_1 and λ_2 , since if λ_1 and λ_2 do not exist then there are two unbounded edges (semilines), one of P and one of Q , that are parallel and either of them determines a line that separates P from Q . Let the point where λ_1 and λ_2 cross be r . Let the point of $P \cap \lambda_1$ closest to r be p_1 and the point of $P \cap \lambda_2$ closest to r be p_2 . Similarly define q_1 and q_2 with respect to Q . It cannot be that both $p_1 = p_2$ and $q_1 = q_2$ so assume, without loss of generality, that $q_1 \neq q_2$. Then there is at least one edge, say e_Q , of Q entirely visible from r . We now show that e_Q separates P from Q . Clearly $L(e_Q)$ intersects λ_1 at a point at least as far from p_1 as is r and similarly intersects λ_2 at a point at least as far from p_2 as is r . Therefore $L(e_Q)$ must separate P from Q since the line parallel to $L(e_Q)$ that intersects r must separate P from Q (clearly, any line that intersects r must either intersect both of P and Q or else separate P from Q). See diagram 2.4. ♦

Edge e_Q separates P from Q

Figure 2.4

Proof of Theorem 2.4. Let $F = (S_1, \dots, S_n)$ be a collection of n disjoint convex sets and let π be as in Theorem 2.2. We denote the interior of convex set S as $\text{int}(S)$. For every element $S_i \in F$ let P_i be the face of π containing S_i . Construct a bipartite graph $G(F, \pi)$ with one vertex $v(m)$ for every line segment R_m of π , $1 \leq m \leq k$, $k \leq 3n-6$; and one vertex $s(i)$ for every convex set S_i of F . The edges of $G(F, \pi)$ are defined such that a vertex $v(m)$ is adjacent to a vertex $s(i)$ if the line $L(R_m)$ does not intersect $\text{int}(S_i)$. For example, for the collection F and partition π illustrated in Figure 2.2, $v(1)$ is adjacent in $G(F, \pi)$ to vertices $s(1)$, $s(2)$, $s(5)$, $s(6)$ and $s(7)$.

We now determine a lower bound for the number of edges in $G(F, \pi)$. We show that, for any pair of elements $S_i, S_j \in F$, there is at least one edge in $G(F, \pi)$. Consider any pair of elements $S_i, S_j \in F$ and the polygonal faces P_i and P_j of π containing them. If P_i and P_j are disjoint then, by Lemma 2.5 there is at least one edge R_r (say of P_i) which separates P_i from P_j . Therefore $L(R_r)$ does not intersect $\text{int}(P_j)$ and $v(r)$ is

adjacent to $s(j)$ in $G(F, \pi)$. If, instead, P_i and P_j share an edge R_t of π then $L(R_t)$ separates S_i from S_j and thus $v(t)$ is adjacent to both $s(i)$ and $s(j)$ in $G(F, \pi)$. We have counted none of the edges of $G(F, \pi)$ more than once. Therefore $G(F, \pi)$ contains at least $\binom{n}{2}$ edges.

By Theorem 2.2, any element R_m of π is on the boundary of faces, say P_{m+} and P_{m-} of π , containing elements of F denoted by S_{m+} and S_{m-} , respectively. Since π has k segments or semilines, $k \leq 3n-6$, there is a vertex $v(m)$ of $G(F, \pi)$ with degree at least $\binom{n}{2}/k \geq \binom{n}{2}/(3n-6) > \frac{n}{6}$. Therefore $L(R_m)$ does not intersect at least $\lceil \frac{n}{6} \rceil$ elements of F . In the worst case, half of them lie on one side of $L(R_m)$ and the remainder lie on the opposite side of $L(R_m)$. In any case, $L(R_m)$ separates one of S_{m+} or S_{m-} from at least $\lceil \frac{n}{12} \rceil$ elements of F .

♦

It is natural at this point to consider if we can determine an upper bound smaller than the obvious upper bound of $\frac{n*(n-1)}{2}$ for convex sets in dimensions higher than two. For $n \geq 3$ the answer is no. This follows from the existence of families of neighborly polytopes, [Grün], that is families of polytopes such that for any pair of them their boundaries have an $n-1$ -facet in common.

3 Separating Circles and Line Segments on the Plane.

In this chapter we study the cases in which the convex sets are all circles or all line segments.

We first study the case that the convex sets are all circles. It is shown that there are collections of n circles C such that no line simultaneously separates any pair of circles of C from any other pair of circles of C . However, it is also shown that there is always one circle out of any collection of n circles C that can be separated from a subcollection of C containing at least $\lfloor \frac{n}{3} \rfloor - 2$ circles. This is a new result which improves on the lower bound of $\lceil \frac{n}{4} \rceil - 1$ circles that can always be separated from one shown in [CCUZ]. The new result is achieved by making modifications to the original result given in [CCUZ].

We then study the case that the convex sets are all line segments. We show that there also exist collections of segments L such that no line simultaneously separates any pair of segments of L from any other pair of line segments of L . We also produce a collection of n line segments J such that no line segment of L can be separated from $\lfloor \frac{n}{3} \rfloor$ elements of L . We show, however, that for any collection of n line segments L that there is a line segment of L that can be separated from at least $\lceil \frac{n-1}{4} \rceil + 1$ elements of L .

Also, when the line segments of L can be extended to semilines so that they remain pairwise disjoint, we show that there is a line that separates one segment of L from at least $\lceil \frac{n}{3} \rceil$ elements of L .

3.1 Upper Bounds on Separating Circles

We now show that for some collections of n circles it is not possible to find two pairs of circles such that one pair can be separated by a line from the other pair. We describe a configuration H_n of n circles in which no pair C_i, C_j in H_n can be simultaneously separated by a single line λ from any other pair C_k, C_l in H_n . Using H_n we construct another example that shows that occasionally one cannot separate one circle from $\lceil \frac{n}{2} \rceil$ circles out of a collection of n circles. Both of these results are from [CCUZ].

Let $S_1 > S_2 > \dots > S_n$ be n different slopes such that $0 \leq S_i \leq \epsilon$, with ϵ small enough. Let H_n consist of n circles defined recursively as follows:

- a) C_1 is any circle in E^2 .
- b) C_{i+1} is a circle tangent to C_i such that the slope of the line that separates them is S_i .
- c) C_{i+1} is large enough such that any line λ separating C_j from C_{i+1} , $1 \leq j < i$ has $\text{slope}(\lambda)$ contained in the interval $(S_i - \delta, S_i + \delta)$, $\delta > 0$, δ much smaller than ϵ .
Observe that $\text{slope}(\lambda)$ is contained in the interval $(-\delta, \epsilon + \delta)$ since $0 \leq S_i \leq \epsilon$.

Moreover, if δ is small enough, C_{i+1} can be chosen such that:

- d) Any line separating C_j from C_i , $1 \leq j < i$, intersects C_{i+1} .

See Figure 3.1. Notice that, in H_n , C_i can be separated from $C_1, \dots, C_{i-1}$, $1 \leq i \leq n$, and that C_i cannot be separated from any pair C_k, C_l , $i < k < l$.

Lemma 3.1. There are no different pairs of circles $\{C_i, C_j\}$ and $\{C_k, C_l\}$ in H_n such that there is a line separating $\{C_i, C_j\}$ from $\{C_k, C_l\}$.

Proof. We assume, without loss of generality, that i is the smallest of i, j, k , and l and that $k < l$. It now follows from (d) that any line separating C_i from C_k must intersect C_l .

◆

The first three circles of H_n
Figure 3.1

Lemma 3.2. There are collections of $2n$ circles C such that no circle of C can be separated from any subcollection of C containing n circles.

Proof. We construct a collection of circles G_n containing $2n$ circles which satisfies the lemma. Construct G_n as follows: Let $H_n' = \{C_1', \dots, C_n'\}$ be the collection of circles H_n (as described above) after being reflected along the x -axis and then translated in the north-west direction until all the lines separating C_i from C_j intersect only C_n' in H_n' , $1 \leq i < j \leq n-1$, and all lines separating C_k' from C_l' intersect only C_n in H_n , $1 \leq i < j \leq n-1$. Now let $G_n = H_n \cup H_n'$. See Figure 3.2.

Any line separating any two elements C_i, C_j in H_n leaves at most $C_1, \dots, C_i$ on one side and $C_1', \dots, C_{n-1}'$ on the other. Therefore G_n is a collection of $2n$ circles and none of them can be separated from any subset of circles in G_n containing more than n circles.

♦

Configuration H_n . No two circles can be separated from two other circles

Figure 3.2

3.2 Separating one Circle from Many Circles.

We have established an upper bound on the number of circles that can always be separated from one in any collection of n circles. We now establish a lower bound. This result is an improvement on the original lower bound given in [CCUZ].

Theorem 3.3. In any collection of n disjoint circles C there is always one that can be separated from a subcollection of C containing at least $\lfloor \frac{n}{3} \rfloor - 2$ circles.

We will need several lemmas and definitions to prove this theorem. We denote the radius of any circle c by $radius(c)$. We say that a set of n disjoint circles $C = \{C_1, \dots, C_n\}$ is *axis-bound* if every circle of C intersects the x -axis and the y -axis but no circle of C contains the origin. The circles of C are assumed to be labeled $C_1, \dots, C_n$ such that the distance from the origin to C_i is smaller than or equal the distance from the origin to C_j , $1 \leq i < j \leq n$. For any quadrant q we further say that C is *axis-bound in q* or simply *in q* if the center of every circle of C is in quadrant q . To prove Theorem 3.3 we first need to prove a number of lemmas that show that given a set C of axis-bound circles we can separate one the circles of C from most of the remaining circles of C . We present the first of these now.

Lemma 3.4. Let $C = \{C_1, \dots, C_m\}$ be m disjoint circles axis-bound in some quadrant, say the positive quadrant. Then any line λ that separates C_m from C_{m-1} also separates C_m from every other element of C .

Proof. Assume for some circle C_i , $1 \leq i \leq m-2$ that $C_i \cap \lambda \neq \emptyset$. Then $C_i \cap \lambda$ must be below the x -axis or left of the y -axis since λ separates C_m from C_{m-1} . Let δ be the line that intersects the x -axis at the point p tangent to C_i , and furthest from the origin. δ must have positive slope since the center of C_i is in the positive quadrant. Also, λ intersects the x -axis at a point to the right of p so it must have positive slope to intersect C_i below the x -axis. Similarly λ must have positive slope to intersect C_i to the left of the y -axis. But λ clearly has negative slope and so it cannot intersect C_i . Therefore λ separates C_m from C_i .

We now prove three lemmas that establish relationships between the radii of circles that are axis-bound. These lemmas are key to proving the remaining lemmas on separating axis-bound circles.

Lemma 3.5. Let C, D be two disjoint circles that do not contain the origin, intersect both the x -axis and the y -axis, and have their centers in the same quadrant, say the positive quadrant. Assume that C is closer to the origin than D . Then $\text{radius}(D) > 4 * \text{radius}(C)$.

Proof. Let C' be a circle which has the same radius as C , does not intersect D , has the origin on its boundary, and has its center on the x -axis or the y -axis. The existence of C' is easily shown by construction. We assume without loss of generality that the center of C' is on the y -axis. Let D' be the circle which has its center in the positive quadrant and is tangent to all of C' , the x -axis, and the y -axis. Now the center of C' is $(0, \text{radius}(C))$ and the center of D' is $(\text{radius}(D'), \text{radius}(D'))$. It is straightforward to show under these conditions that $\text{radius}(D') = 4 * \text{radius}(C)$. But clearly $\text{radius}(D') > \text{radius}(D)$ so $\text{radius}(D) > 4 * \text{radius}(C)$.

♦

For any set of n circles C axis-bound to the same quadrant we denote an arbitrarily chosen line which separates C_n from C_{n-1} as $\lambda(C)$.

Lemma 3.6. Let $C = \{C_1, \dots, C_m\}$ and $D = \{D_1, \dots, D_n\}$ be two sets of m and n disjoint circles respectively such that they are axis-bound to non-adjacent quadrants, say the positive quadrant and the negative quadrant respectively. Assume that $\lambda(D)$ intersects a circle C_i of $C \setminus \{C_m\}$. Then the radius of C_i is larger than the radius of any circle of $D \setminus \{D_n\}$.

$$\text{radius}(C_i) > \text{radius}(D_{n-1})$$

Figure 3.3

Proof. We must show that $\text{radius}(D_{n-1}) < \text{radius}(C_i)$ since D_{n-1} is the largest circle of D after D_n . We assume, without loss of generality, that $C_i \cap \lambda(D)$ is in the quadrant below the positive quadrant. Let d_i be the distance from $C_i \cap \lambda(D)$ to the x -axis and d_y be the distance from the intersection of $\lambda(D)$ and the y -axis to the origin. Then, since $\lambda(D)$ has negative slope, $d_i > d_y$. But clearly $d_i < \text{radius}(C_i)$ and $d_y > \text{radius}(D_{n-1})$. Therefore $\text{radius}(C_i) > \text{radius}(D_{n-1})$. See Figure 3.3.

◆

Corollary 3.7. If, instead, D is in a quadrant adjacent to the positive quadrant then twice the radius of C_i is larger than the radius of any circle of $D \setminus \{D_n\}$.

Proof. Assume, without loss of generality, that D is in the quadrant directly below the positive quadrant. We must show that $\text{radius}(D_{n-1}) < 2 * \text{radius}(C_i)$ since D_{n-1} is the largest circle of D after D_n . Let d_i be the distance from $C_i \cap \lambda(D)$ to the y -axis and d_h be the distance from the intersection of $\lambda(D)$ and the x -axis to the origin. Then, since $\lambda(D)$ has positive slope, $d_i > d_y$. But clearly $d_i < 2 * \text{radius}(C_i)$ and $d_h > \text{radius}(D_{n-1})$. Therefore $2 * \text{radius}(C_i) > \text{radius}(D_{n-1})$.

♦

We now prove the remaining lemmas on separating axis-bound circles. The following lemma is key to proving the rest.

Lemma 3.8. Let $C = \{C_1, \dots, C_m\}$ and $D = \{D_1, \dots, D_n\}$ be two sets of m and n disjoint circles respectively such that they are axis-bound to different quadrants. Let $H = (C \setminus \{C_m\}) \cup (D \setminus \{D_n\})$. Then either $\lambda(C)$ separates C_m from H or $\lambda(D)$ separates D_n from H (or both).

Proof. By Lemma 3.4 $\lambda(C)$ separates C_m from $(C \setminus \{C_m\})$ and $\lambda(D)$ separates D_n from $D \setminus \{D_n\}$. We assume, without loss of generality, that C is axis-bound to the positive quadrant. We also assume that $\lambda(D)$ does not separate D_n from $C \setminus \{C_m\}$ since otherwise $\lambda(D)$ would satisfy this lemma. Then $\lambda(D)$ must intersect at least one element C_i of $C \setminus \{C_m\}$. Now there can be no element D_j of $D \setminus \{D_n\}$ that intersects $\lambda(C)$. To show this we assume D_j exists and produce a contradiction.

Consider first that D is axis-bound to a quadrant adjacent to the positive quadrant, say the quadrant directly below the positive quadrant. Let p be point of $D_j \cap \lambda(C)$. Then p

must be in one of the four quadrants. The positive quadrant and the quadrant directly below the positive quadrant cannot contain p because this would require that D_j intersect either C_i or $\lambda(D)$. The negative quadrant cannot contain p since no part of $\lambda(C)$ is in the negative quadrant. Finally, p is not in the quadrant directly to the left of the positive quadrant since no part of D_j is in that quadrant. Therefore p does not exist and so C and D are not in adjacent quadrants. See Figure 3.4.

C and D cannot be in adjacent quadrants

Figure 3.4

So D must be axis-bound to the negative quadrant. Then, since $\lambda(D)$ intersects C_i , $\text{radius}(C_i)$ is larger than the radius of any circle of $D/\{D_n\}$ by Lemma 3.6. Similarly,

since $\lambda(C)$ intersects D_i , $\text{radius}(D_i)$ is larger than the radius of any circle of $C/\{C_m\}$; a contradiction.

Thus no matter which of the three possible quadrants is chosen for D either $\lambda(C)$ separates C_m from H or $\lambda(D)$ separates D_n from H (or both).

Let $A = \{A_1, \dots, A_r\}$, $B = \{B_1, \dots, B_s\}$, $C = \{C_1, \dots, C_t\}$, and $D = \{D_1, \dots, D_u\}$ be four sets of r, s, t , and u disjoint circles respectively, which are axis-bound to different quadrants. Let $n = r + s + t + u$. For any two elements of A, B, C, D , say A, B , we say that $B > A$ if $\lambda(A)$ separates A_r from $B/\{B_s\}$. We also say that $A \gg B$ if $\lambda(A)$ does not separate A_r from $B/\{B_s\}$ and that $A \ggg B$ if $\lambda(A)$ does not separate A_r from $B/\{B_s, B_{s-1}\}$. Note that by Lemma 3.8 the relations " $\gg$ " and " $\ggg$ " are anti-symmetric. We now have three lemmas relating to A, B, C , and D .

Lemma 3.9. For any three of A, B, C , and D in any order, say A, B , and C if $A \gg B \ggg C$ (or $A \ggg B \ggg C$) then $A > C$.

Proof. We assume $C \gg A$ and produce a contradiction. We will prove only one case: A is in the positive quadrant, B is in the negative quadrant, and C is in the quadrant to the left of the positive quadrant since all other cases are symmetric to this one (once we assume $C \gg A$). Let A_i be any circle of $A/\{A_r\}$ intersected by $\lambda(C)$, B_j be any circle of $B/\{B_s\}$ intersected by $\lambda(A)$, and C_k be any circle of $C/\{C_t, C_{t-1}\}$ intersected by $\lambda(B)$.

To start we know that $\text{radius}(B_j) > \text{radius}(A_i)$ because $\lambda(A)$ intersects B_j and by Lemma 3.6. Then, since $\lambda(B)$ intersects C_k and since B and C are axis-bound to adjacent quadrants, $2 * \text{radius}(C_k) > \text{radius}(B_j) > \text{radius}(A_i)$ by Corollary 3.7. Also, $\text{radius}(C_{m-1}) > 4 * \text{radius}(C_k) > 2 * \text{radius}(A_i)$ by Lemma 3.5 and the preceding

sentence. Since A and C are in adjacent quadrants we have that $2 * \text{radius}(A_i) > \text{radius}(C_{m-1})$ by Corollary 3.7. Therefore $2 * \text{radius}(A_i) > \text{radius}(C_{m-1}) > 2 * \text{radius}(A_i)$; a contradiction. Therefore, if $A \gg B \gg C$ then $A > C$.

In fact this lemma can be strengthened to prove that $A \gg B \gg C \Rightarrow A > C$ as can be seen from Figure 3.5. This stronger result is valid because we cannot place C_k so that it intersects both axes but does not intersect A_i and still have $\lambda(C)$ intersect A_i (even if we let $C_k = C_{t-1}$). However, the improvement to Theorem 3.3 that results from this stronger result is minimal and we would also need to make a similar improvement to the following lemma.

Lemma 3.10. For any ordering of A, B, C , and D , say A, B, C , and D , if $A \gg B \gg C \gg D$ then $A > D$.

We will only outline the prove this lemma since its proof so similar to the proof of Lemma 3.9. Let A_i be any circle of $A/\{A_r\}$ intersected by $\lambda(D)$, B_j be any circle of $B/\{B_s, B_{s-1}\}$ intersected by $\lambda(A)$, C_k be any circle of $C/\{C_t, C_{t-1}\}$ intersected by $\lambda(B)$, and D_l be any circle of $D/\{D_u, D_{u-1}\}$ intersected by $\lambda(C)$.

If we consider the ordering A, B, C, D as cyclic then up to symmetry, it is necessary to prove two cases:

- 1) there are exactly two consecutive pairs, say AB and CD , that are in adjacent quadrants.
- 2) All four of the consecutive pairs of the ordering are in adjacent quadrants.

In either case we show, using the same lemmas and the same corollary used in the previous proof, that if $A > D$ then $\text{radius}(A_i) < \text{radius}(B_{s-1}) < \text{radius}(C_{t-1}) < \text{radius}(D_{u-1}) < \text{radius}(A_i)$; a contradiction. It may also be that if $A \gg B \gg C \gg D$ then $A > D$. The improvement to Theorem 3.3 that would result, however, is minimal.

Figure 3.5

We are almost home. The previous two lemmas are used to prove the following lemma, which is the one needed to prove Theorem 3.3. Let $ABCD$ denote the set $\{A_1, \dots, A_r, B_1, \dots, B_s, C_1, \dots, C_t, D_1, \dots, D_u\}$.

Lemma 3.11. One of the lines $\lambda(A)$, $\lambda(B)$, $\lambda(C)$, and $\lambda(D)$ separates one element of $ABCD$ from at least $|ABCD| - 5$ elements of $ABCD$.

Proof. We assume the lemma is false and produce a contradiction. Consider the graph G constructed as follows: Let the vertices of G correspond to sets A , B , C , and D . Let there be a directed edge from the vertex corresponding to A to the vertex corresponding to B of G if $B > A$. (i.e. $\lambda(A)$ separates A_r from every circle of $B \setminus (B_s)$). Similarly define all the edges of G . We know by Lemma 3.8 that G has at least one edge between any two vertices. If there are two edges between any two vertices of G we arbitrarily discard one of them. (It will be evident from the remainder of the proof that we can safely discard these edges.) Then G must have six edges in all and there are four possible graphs, say G_1 , G_2 , G_3 , and G_4 , for G after eliminating symmetries. See Figure 3.6. In graphs G_1 and G_2 it is clear from the definition of the operator ' $>$ ' that $\lambda(A)$ separates A_r from $|ABCD| - 4$ elements of $ABCD$. The analysis for G_4 is similar to that of G_3 so now we only need consider G_3 . The remainder of the proof applies only to G_3 .

Since $B > A$ and $C > A$ in graph G_3 it must be that $A \gg D$ or else $\lambda(A)$ separates A_r from $|ABCD| - (4 \text{ or } 5)$ elements of $ABCD$. ($\lambda(A)$ would fail to separate A_r only from the elements B_s, C_t, D_u , and perhaps D_{u-1} of $ABCD$.) Similarly $D \gg C$ in G_3 .

We know that $A > C$ since $A \gg D \gg C$ and by Lemma 3.9. Then $C \gg B$ or else $\lambda(C)$ separates C_t from $|ABCD| - (4 \text{ or } 5)$ elements of $ABCD$ since $D > C$, $A > C$ and by Lemma 3.9. This, in turn, implies that $D > B$ since $D \gg C \gg B$. We similarly show that $B \gg A$. We now have that $A \gg D \gg C \gg B \gg A$. But this is impossible by Lemma 3.10. Therefore one of the lines $\lambda(A)$, $\lambda(B)$, $\lambda(C)$, and $\lambda(D)$ separates one element of $ABCD$ from at least $|ABCD| - 5$ elements of $ABCD$. ♦

The four possible ways $A, B, C,$ and D can be related after eliminating symmetries.

Figure 3.6

Lemma 3.4 is the lemma we need to finally prove Theorem 3.3.

Proof of Theorem 3.3. Let $\{C_1, \dots, C_n\}$ be the set of circles of C . Start by sweeping a line λ_1 , left to right and parallel to the y -axis until one element $C_j \in C$ is to the left of λ_1 . Then sweep a line λ_2 , from bottom to top and parallel to the x -axis until one circle, say C_2 , is below λ_2 . Let n_1 be the number of circles to the right of λ_1 and n_2 be the number of circles above λ_2 . If $n_1 \geq \lfloor \frac{n}{3} \rfloor - 2$, or $n_2 \geq \lfloor \frac{n}{3} \rfloor - 2$ then the result holds. Suppose then that $n_1, n_2 < \lfloor \frac{n}{3} \rfloor - 2$. Then there is a subset H of C with $n -$

(n_1+n_2+2) circles that intersect both of λ_1 and λ_2 and at most one of them, say C_3 , contains the intersection point of λ_1 and λ_2 . Therefore $|H(C_3)| \geq \lfloor \frac{n}{3} \rfloor + 3$.

Now consider λ_1 and λ_2 as the coordinate axes. Then, by Lemma 3.5, some line separates one element of $H(C_3)$ from at least $\lfloor \frac{n}{3} \rfloor - 2$ elements of $H(C_3)$. Therefore there is a line that separates one element of C from at least $\lfloor \frac{n}{3} \rfloor - 2$ elements of C .

◆

If the minor improvements suggested to the proofs of Lemma 3.9 and Lemma 3.10 are true then this result can be improved to $\lfloor \frac{n+1}{3} \rfloor - 2$. Note also that for some small values of n the lower bound of $\lceil \frac{n}{4} \rceil - 1$ circles that can always be separated from one shown in [CCUZ] is slightly better than this new result.

We have seen from Lemma 3.2 that we cannot generally find a line that separates two circles from two circles out of a collection of n circles. The situation is entirely different, however, if we also require that the circles be congruent. It is then possible to separate $O(\frac{n}{2})$ circles from $O(\frac{n}{2})$ circles. This is a result of the following theorem given in [AKP].

Theorem 3.12. There exists a constant $c > 0$ such that if C is any collection of n congruent disjoint circles then there is a line that separates at least $\frac{n}{2} - c \sqrt{n} \sqrt{\log(n)}$ elements of C from at least $\frac{n}{2} - c \sqrt{n} \sqrt{\log(n)}$ elements of C .

This is a particularly strong result considering Lemma 3.2. The proof is quite long and difficult, however, and we choose not to present it here. We refer the interested reader to the original paper [AKP].

3.3 Upper Bounds on Separating Line Segments

We now consider the case that the convex sets are closed and bounded line segments. Except where otherwise mentioned all the remaining results of this chapter are from [CCUZ]. We first determine upper bounds on separating line segments.

We first show that it is not always possible to separate any pair of line segments from any other pair of segments out of an arbitrary collection of line segments. We describe a configuration T_n of n line segments in which no pair S_i, S_j in T_n can be simultaneously separated by a single line λ from any other pair S_k, S_l in T_n . Using T_n we construct another configuration J_n that shows that occasionally one cannot separate one segment of J_n from any subset of $\lceil \frac{n}{3} \rceil$ segments of J_n .

The configuration T_n is a slightly modified version of a configuration due to K.P. Villanger. Villanger's version was used in [Tver] to prove this same result. This was the first proof that it is not always possible to separate two convex sets from two convex sets out of a collection of n convex sets. We use T_n here, however, since we also need it to construct J_n . The configurations T_n and J_n are due to [CCUZ].

The family of line segments $T_n = S_1, \dots, S_n$ is constructed as follows: Let the segments $S'_1, \dots, S'_n$ have slopes $0 = \text{slope}(S'_1) < \text{slope}(S'_2) < \dots < \text{slope}(S'_n) = \delta < \frac{\pi}{2}$, respectively; and such that the left endpoint of S'_{i+1} is an interior point of S'_i , $2 \leq i < n$. Also, the left endpoint of S'_i is within distance ϵ of the left endpoint of S'_1 , $2 \leq i \leq n$. Finally, let S_i be the segment obtained by shortening the left end of S'_i by a sufficiently small amount.

No line separates S_k from $\{S_i, S_j\}$

Figure 3.7

Theorem 3.13. No two elements of T_n can be separated by a line from any other two elements of T_n [Tver].

Proof. We observe that for any $S_i, S_j, S_k \in T_n$, $1 \leq i < j < k \leq n$, that S_k intersects the convex closure of $S_i \cup S_j$ and therefore no line separates S_k from $\{S_i, S_j\}$. The lemma now follows immediately. See Figure 3.7.

◆

We now construct J_n from T_n . J_n consists of three copies $\Gamma_0 = \{S_{0,1}, \dots, S_{0,n}\}$, $\Gamma_1 = \{S_{1,1}, \dots, S_{1,n}\}$, and $\Gamma_2 = \{S_{2,1}, \dots, S_{2,n}\}$ of T_n placed around a triangle with vertices v_0, v_1, v_2 such that $S_{i,1}$ lies on the line determined by v_i and v_{i+1} , and v_i is sufficiently close to the left endpoint of $S_{i,1}$; addition taken mod 3. See Figure 3.8. The values of ϵ and δ are chosen small enough that any element of T_i , when extended to be a whole line, intersects all the elements of T_{i+1} , $0 \leq i \leq 2$; addition taken mod 3.

Configuration J_n

Figure 3.8

Theorem 3.14. No element in J_n can be separated by a line from more than $n+1$ elements in J_n . [CCUZ].

Proof. Every line that separates an element of Γ_i from an element of Γ_{i+1} must intersect every segment in Γ_{i+2} , $0 \leq i \leq 2$; addition taken mod 3. Therefore no element of Γ_i can be simultaneously separated by a single line from one element in Γ_{i+1} and one element in Γ_{i+2} . Also any line λ that separates two elements $S_{i,r}$ and $S_{i,s}$, $1 \leq r < s \leq n$, must intersect every element of $\{S_{i,t}, 1 \leq t < s\}$ and these are the elements of Γ_i on the same side of λ as Γ_{i+2} . Therefore the best we can do is a line that simultaneously separates one element of Γ_i from one other element of Γ_i and from every element of Γ_{i+2} .

♦

This result also represents the best upper bound we know on the number of convex sets that can be separated from one out of a collection of n convex sets.

3.4 Separating one Line Segment from Many.

We have established an upper bound on the number of line segments that can always be separated from one in any collection of line segments. We now establish a lower bound. Let $L(S)$ denote the line determined by line segment S .

Theorem 3.15. In any collection of n disjoint closed and bounded line segments F there is always one that can be separated from a subcollection of F containing at least $\lceil \frac{n-1}{4} \rceil + 1$ segments.

Proof. Let $S_1, \dots, S_n$ be the line segments of F . Construct a graph G as follows: Let the vertices of G correspond to the elements of F . For any two vertices v_i, v_j of G there is a directed edge $v_i \rightarrow v_j$ from v_i to v_j if $L(S_i)$ intersects S_j . Clearly if $v_i \rightarrow v_j$ is in G then $v_j \rightarrow v_i$ is not in G . It follows that there is a vertex v_k of G with out-degree of at most $\lfloor \frac{n-1}{2} \rfloor$. Then $L(S_k)$ intersects at most $\lfloor \frac{n-1}{2} \rfloor$ elements of F and thus does not intersect at least $\lceil \frac{n-1}{2} \rceil$ elements of F . Then there is a subcollection F' of F containing at least $\lceil \frac{n-1}{2} \rceil$ line segments on one side of $L(S_k)$ and a collection F'' containing $\lceil \frac{n-1}{2} \rceil - |F'|$ elements of F on the other side of $L(S_k)$. Let λ be a line parallel to $L(S_k)$, arbitrarily close to $L(S_k)$. If $F'' = \emptyset$ then choose λ such that it separates S_k from F' . Then λ satisfies this lemma. Otherwise choose λ such that S_k and F' are on the same side of it. Then λ separates the elements of F'' from $S_k \cup F'$. But $|S_k \cup F'| \geq \lceil \frac{n-1}{4} \rceil + 1$.

♦

If our line segments can be extended to be semilines so that they remain pairwise disjoint then this result can be improved.

Theorem 3.16. Let $F = \{S_1, \dots, S_n\}$ be a collection of n disjoint line segments, $n \geq 4$, such that they can be extended to form a collection of disjoint semilines. Then there is a line λ that separates an element S_i of F from a subcollection of F containing at least $\lfloor \frac{n}{3} \rfloor + 1$ elements [CCUZ].

Proof. If there is an element $S_i \in F$ such that $L(S_i)$ does not intersect any other element of F , then S_i can be separated from a subcollection of F containing at least $\lfloor \frac{n-1}{2} \rfloor$ elements of F . Suppose then that each $L(S_i)$ $1 \leq i \leq n$, intersects at least one other element of F . Extend the elements of F to be semilines so that they remain pairwise disjoint and then extend them (on the other side) as much as possible until a collection $F' = \{S_1', \dots, S_n'\}$ of semilines is obtained such that:

- 1) The endpoint of every element of F' lies on an interior point of another element of F' .
- 2) No two elements of F' cross each other. See Figure 3.9.

We say that S_i' hits S_j' if the endpoint of S_i' lies on S_j' , $S_i', S_j' \in F'$. For example, in Figure 3.9, S_1' hits S_4' . Since the endpoint of every $S_i' \in F'$ is an interior point of another element of F' , there must be a cyclic sequence of elements, say $S_1', \dots, S_j'$, $j \leq n$, such that L_{i+1}' hits L_i' , $i = 1, \dots, j$; addition taken mod j .

For the case when $j = n$ it is easily shown by constructing a graph similar to the one in the previous theorem (and considering the in-degree of the vertices) that $\lfloor \frac{n-1}{2} \rfloor$ elements of F can be always be separated from one. We therefore assume that $j < n$.

Extending semilines until they touch.

Figure 3.9

For every $i = 1, \dots, j$ let T_i be the subset of F consisting of S_i' together with all the elements of F contained in the region bounded by S_i' and S_{i-1}' ; addition taken mod j . For the example in Figure 3.9, $T_1 = \{S_1', S_5', S_6'\}$, $T_2 = \{S_2'\}$, $T_3 = \{S_3'\}$, and $T_4 = \{S_4'\}$. Since $j < n$ there is a k such that T_k contains more than one element. Suppose, without loss of generality, that $k = 1$ (as in Figure 3.9).

Let i be the smallest index such that $L(S_1')$ intersects S_i' . Let $A = T_2 \cup \dots \cup T_{i-1}$, $B = T_i$, and $C = T_{i+1} \cup \dots \cup T_j \cup T_1$. Then clearly A is separable from S_1 , B is separable from S_{i-1}' , and C is separable from S_i' ; see Figure 3.10.

However, since $A \cup B \cup C = F$, at least one of A , B , and C has at least $\lceil \frac{n}{3} \rceil$ elements. Moreover, if not all their cardinalities are the same, then at least one of them has $\lfloor \frac{n}{3} \rfloor + 1$ elements and the result is proved. So assume then that A , B , and C all have the same cardinality. But then, since $|T_1| > 1$, there must be a $S_a' \in T_1, - \{S_1'\}$ and it is clear that S_a is separable from $A \cup \{S_1'\}$.

◆

Decomposition of F' into three subsets

Figure 3.10

The segments in the example J_m may be extended to semilines in such a way that they remain pairwise disjoint. This shows that the bound on this theorem is tight. This is the first instance we know of in which a tight bound has been determined for the separation of one element from many. In Chapter four we will see that this can also be done for point sets in the plane.

It is worth noting and easy to verify that results of this chapter remain valid even if the convex set are permitted to be adjacent but must still have pairwise disjoint relative interiors.

4 Separating Points

In this chapter we study the problem of separating a set of n points P in general position in dimension d by a set of hyperplanes of dimension $d - 1$. We denote the d -dimensional Euclidean space by E^d . For any two points $a, b \in P$, we denote the open line segment joining them $\text{seg}(a, b)$ and the closed line segment joining them $\text{seg}[a, b]$. We say that a set of points P in E^d is *separated* by a set of hyperplanes H in E^d if for every pair of points $a, b \in P$, $\text{seg}(a, b)$ is intersected by at least one hyperplane h of H . We further say that h *separates* a from b . See Figure 4.1.

A set of ten points separated by four lines
Figure 4.1.

Let $f(n, d)$ be the minimum number of hyperplanes always sufficient and occasionally necessary to separate n points in E^d where $n \geq d$. We show that $\lceil \frac{n-1}{d} \rceil \leq f(n, d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d odd and $\lceil \frac{n}{d} \rceil \leq f(n, d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d

even. For dimension three we improve this result slightly. From now on we assume that $n \geq d$. Theorem 4.7 will show that if $n < d$ then $f(n,d) = \lceil \log(n) \rceil$.

The upper bounds of $f(n,d)$ are determined by construction: we present algorithms that generate a solution within these upper bounds. The algorithms are efficient, ranging in complexity from $O(n \log(n))$ to $O(n^2)$. The $O(n^2)$ algorithm is very easy to implement and very fast in practice, having an expected complexity of $O(n \log(n))$.

By 'general position' in E^d we mean that no $d + 1$ points are in the same hyperplane in E^d . This restriction is important since if the points can be located arbitrarily then we can trivially show that $n - 1$ hyperplanes are needed simply by placing n points in a straight line. From now on all points sets considered here will be assumed to be in general position.

This chapter is structured as follows: First we establish the lower bounds for $f(n,d)$. Next we present an $O(n^2)$ algorithm to separate point sets in E^d ; in the process the upper bounds on $f(n,d)$ are established. We then consider some possible improvements to the algorithms. First we consider some ways of reducing the upper bounds even further for $d \geq 2$. We then show that in E^2 we can reduce the time complexity of our algorithm to $O(n \log(n))$ and in E^3 we can reduce it to $O\left(\frac{n^{3/2} \log^3(n)}{\log^*(n)}\right)$. Finally we develop a data structure for storing the lines or hyperplanes so that retrieving from the data structure a line or hyperplane that separates two given points can be found in worst case time $O(n)$ and expected case time $O(\log(n))$. If $d = 2$ we show that we can construct an even more efficient data structure that allows us to retrieve a line that separates two given points in worst case time $O(\log(n))$ and expected case time $O(1)$.

4.1 Determination of Lower Bounds for $f(n,d)$

Determining a lower bound for $f(n,2)$ is straightforward. Let P be the set of vertices of a convex polygon Q with n vertices and let S be any set of lines that separates P . Then every edge of Q has to be intersected by at least one element of S , and since no line intersects more than two of the edges of any convex polygon, $|S| \geq \lceil \frac{n}{2} \rceil$. Clearly $\lceil \frac{n}{2} \rceil$ will do. See Figure 4.2.

A collection of six points for which three lines are needed to separate them

Figure 4.2

To determine a lower bound for $f(n, d)$ consider the moment curve

$$M_d(t) = (t, t^2, \dots, t^d) \quad (4.1)$$

in E^d . Let $\Psi_n(d) = \{M_d(1), \dots, M_d(n)\}$ on $M(t)$ be a set of n points on M_d . We determine how many times a hyperplane can intersect the moment curve. A hyperplane in E^d can be represented by an equation of the form:

$$a_d x_d + a_{d-1} x_{d-1} + \dots + a_1 x_1 + a_0 = 0 \quad (4.2)$$

Combining (4.2) and (4.1) we get the equation

$$a_d t^d + a_{d-1} t^{d-1} + \dots + a_1 t + a_0 = 0 \quad (4.3)$$

Equation (4.3) is a polynomial of degree d and thus has at most d roots. This proves the following well known result.

Lemma 4.1. Any hyperplane h in E^d intersects the moment curve M_d in at most d points. If h intersects M_d in exactly d points then h does not intersect M_d tangentially.

Using Lemma 4.1 we establish lower bounds for $f(n,d)$.

Theorem 4.2. $f(n,d) \geq \lceil \frac{n-1}{d} \rceil$ for d odd and $f(n,d) \geq \lceil \frac{n}{d} \rceil$ for d even.

Proof. Consider $\Psi_n(d) = \{M_d(1), \dots, M_d(n)\}$. Let $M(i,j)$ be the section of M_d between points $M_d(i)$ and $M_d(j)$, $i < j$. Let e_i denote $\text{seg}[M_d(i), M_d(i+1)]$ and let $E = \{e_i \mid 1 \leq i < n\}$. Let H be any set of hyperplanes that separates $\Psi_n(d)$. Observe that every $e_i \in E$ has to be crossed by some hyperplane of H . See Figure 4.3.

We determine the maximum number of segments of E that any hyperplane $h \in H$ can intersect. Assume h intersects an element e_i of E . Then h separates $M_d(i)$ from $M_d(i+1)$. Therefore, since $M(i,i+1)$ is an arc joining $M_d(i)$ and $M_d(i+1)$, h intersects $M(i,i+1)$. (Note that we need not consider that h intersects $M(i,i+1)$ at points $M_d(i)$ or $M_d(i+1)$ since h is to separate these points.) Now, by Lemma 4.1, h intersects M_d at most d times; therefore h crosses at most d elements of E . Since $|E| = n - 1$, we have that $f(n,d) \geq \lceil \frac{n-1}{d} \rceil$.

If d is even then this result can be improved slightly. We observe that the line segment $f = \text{seg}[M_d(1), M_d(n)]$ must also be intersected by the hyperplane $h \in H$ that separates $M_d(1)$ from $M_d(n)$. If h crosses $M(1,n)$ d times then each intersection is a

crossing of $M(1,n)$ by Lemma 4.1 and since d is even $M_d(1)$ and $M_d(n)$ lie on the same side of h , contradicting that h separates them. Therefore h crosses $M(1,n)$ at most $d - 1$ times. Therefore, when d is even, $f(n,d) \geq \lceil \frac{n}{d} \rceil$.

♦

Five points on the moment curve

Figure 4.3

4.2 A Simple Point Separation Algorithm

We now develop an algorithm that, given any collection P of n points in E^d , produces a set of $\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ hyperplanes that separate P . Therefore $f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$.

Let P be a point set in E^d and H be any set of hyperplanes in E^d ; a cell CL of H is a connected subset of E^d such that, for any two points $a, b \in CL$, $\text{seg}(ab)$ does not intersect any hyperplane of H . We say that a subset C of P is a *cluster* of H if C is contained in a cell CL determined by some subset of H while $P - C$ is contained in the

complement of CL . A partitioning $S = \{P_1, \dots, P_k\}$ of P into disjoint subsets is called a *cluster set* of (P, H) . If each P_i is a cluster of H , $i=1, \dots, k$; $\text{Point}(S)$ denotes P and $\text{Hyper}(S)$ denotes H . See Figure 4.4. Note that (P, H) does not, in general, determine S .

A possible cluster set for a set of fourteen points and four lines

Figure 4.4.

Note that if all the elements of S are singletons, then H separates P . Let $|C|$ denote the cardinality of a cluster C . We say that C is *separated* if $|C| = 1$ and *crowded* if $|C| \geq 2$. Given a cluster set S_i of (P, H) we say that a hyperplane h (not in H) refines S if there is a crowded cluster C of S such that h splits C into two smaller non-empty clusters C_1, C_2 . We then say that h is a *1-refine* of S_i and denote by $S_{i+1} = (S_i - C) \cup \{C_1, C_2\}$. Where the meaning is clear we simply say that h_i 1-refines C . A 1-refine of S where $S = \{P\}$ may be simply referred to as a 1-refine of P . A single hyperplane h_i that simultaneously splits k crowded clusters of S will be called a *k-refine* of S .

Observation: 4.3. Let $S_0 = \{P\}$ and $H = \{h_1, h_2, \dots, h_{n-1}\}$ be a family of hyperplanes such that h_i is a 1-refine of S_i , $i = 1, \dots, n-1$. Then H separates P .

The above observation gives rise to an easy algorithm to separate an n -point set in E^d with $n-1$ hyperplanes. Although this is an easy algorithm that will not give the number of hyperplanes needed to prove our upper bounds, it will be of importance later on in our complexity analysis.

Algorithm 1-Separate.

Input: An n -point set P in general position in E^d .

Output: A set of hyperplanes that separates P .

- 1) $S := \{P\}$, $H = \emptyset$.
- 2) For $i := 1$ to $n-1$ do
 - 2.1) Choose $p_i, p_j \in F_i$, F_i a cluster of S and a hyperplane h_i that intersects $\text{seg}[p_i, p_j]$.
 - 2.2) $S := (S - F_i) \cup \{C_i, D_i\}$ where C_i, D_i are the two clusters that result from a 1-refine of F_i by h_i . $H = H \cup h_i$
- 3) Return (H) .

Theorem 4.4. Algorithm 1-Separate separates an n point set P in time $O(n^2)$ in the worst case and $O(n \log(n))$ in the expected case regardless of how the points are distributed.

Proof. We show that 1-separate is equivalent to a version of algorithm Quicksort and has the same run time behavior. Quicksort invokes a subalgorithm 'Partition' which partitions a subcollection into two smaller subcollections. We observe that 1-separate and Quicksort are logically equivalent ignoring the internal behavior of 1-refine and Partition.

We need only show that 1-refine partitions the points in a cluster 'equally as well' as Partition partitions a collection of values.

We choose a version of Partition that determines a value x to use to partition a subcollection A by randomly selecting two values $a, b \in A$, $a \leq b$, and choosing x such that $a \leq x \leq b$. Of the values a, b let y be the one which is the furthest from the median of A . Clearly Partition performs worst when $x = y$. However even with this bad choice for x the $O(n \log(n))$ expected run time behavior of Quicksort is unchanged. We also note that the analysis of Quicksort is independent of the distribution of the values being sorted.

Now consider Algorithm 1-refine. Let us assume that 1-refine selects a hyperplane h so as to partition a cluster as unevenly as possible but that it does so before randomly selecting the two points that determine a line segment s that h must intersect. That is, it chooses h such that the 'average unevenness' over all possible choices of points is as bad as possible. First 1-refine must select a set T of $n - 1$ points that along with some point $s_h \in s$ will determine h . Once s is determined 1-refine is free to make as bad a choice for s_h as possible. Let Ψ be any hyperplane that intersects each point of T . For each point p in E^d we associate a value $\omega(p, \Psi)$ which is the angle between the hyperplane Ψ and the hyperplane determined by T and p . Let $\Psi(C)$ be the set of values determined by applying $\omega(p, \Psi)$ to each point p in cluster C . Since the values $\Psi(C)$ are totally ordered (except for possible ties) they can be partitioned by algorithm Partition. Let p_i, p_j be the endpoints of s such that $\omega(p_i, \Psi) \leq \omega(s_h, \Psi) \leq \omega(p_j, \Psi)$. We can also assume that Partition partitions $\Psi(C)$ using value $\omega(s_h, \Psi)$ since choosing p_i, p_j randomly corresponds to choosing $\omega(p_i, \Psi), \omega(p_j, \Psi)$ randomly. That is, for every random pair $p_i, p_j \in C$ there exists exactly one corresponding pair $\omega(p_i, \Psi), \omega(p_j, \Psi) \in \Psi(C)$. Thus Partition will partition $\Psi(C)$ into two parts $C_1(\Psi)$ and $C_2(\Psi)$ where C_1, C_2

are the two clusters produced by the 1-refine of C by h . Therefore 1-refine partitions clusters equally as well as Partition partitions collections. Therefore 1-separate has the same run time behavior as Quicksort. ♦

4.3 Separating Points by Lines on the Plane

We now show that $f(n,2) = \lceil \frac{n}{2} \rceil$ by constructing an algorithm that separates a point set P with $\lceil \frac{n}{2} \rceil$ lines. We first make a crucial observation: given any two pairs of distinct points p_1, p_2 and q_1, q_2 in general position there exists a line that simultaneously intersects $\text{seg}(p_1, p_2)$ and $\text{seg}(q_1, q_2)$. An example is the line determined by the midpoints of these line segments (if one of these midpoints is aligned with the line determined by the other pair, a small perturbation of that line will do). This gives us a simple algorithm for carrying out a 2-refine.

Algorithm 2-refine

Input: A cluster set S and two crowded clusters $C, D \in S$.

Output: 1) A cluster set S' that results from a simultaneous 1-refine of C and of D by a line λ .

2) the line λ .

- 1) Select any four points $c_1, c_2 \in C$ and $d_1, d_2 \in D$.
- 2) $\lambda :=$ the line determined by the midpoints of $\text{seg}(c_1, c_2)$ and $\text{seg}(d_1, d_2)$, (or a small perturbation of it).
- 3) Partition each of C and D into two clusters C_1 and C_2 (D_1 and D_2 resp.) using λ according to which side of the line their elements are on.
- 4) Return($(S - C - D) \cup \{C_1, C_2, D_1, D_2\}, \lambda$).

For example see Figure 4.5 (a).

Consider an n -point set P and a line ∂ such that ∂ splits P into two subsets P_1 , P_2 and $|P_1| = |P_2|$. Let $S = (P_1, P_2)$. We can do simultaneous 1-refines of two clusters of S , one on each side of ∂ , using Algorithm 2-refine. This leads us to our second point separation algorithm. See Figure 4.5 (a).

Algorithm 2-Separate.

Input: A point set P in the plane.

Output: A set of lines L that separates P .

- 1) Partition P into two point sets P_1 and P_2 by a line ∂ such that $|P_1| \leq |P_2| \leq |P_1| + 1$.
- 2) $S := (P_1, P_2)$, $H = \{\partial\}$, $i := 1$.
- 3) While there is a crowded cluster C_i to the left of ∂ and a crowded cluster D_i to the right of ∂ find a line λ_i that simultaneously splits C_i and D_i . This is accomplished using 2-refine. $H := H \cup \lambda_i$, $i := i + 1$.
- 4) If $|P_1| < |P_2|$ then there is one crowded cluster remaining (on the side of ∂ containing P_2). Add an extra line λ_i that splits it into two, and set $H := H \cup \lambda_i$.
- 5) Return H .

See Figure 4.5.

♦

The sequence of the first three 2-splits ((a), (b), and (c)) of P
and the final separating set of lines of P (d).

Figure 4.5

Theorem 4.5. Algorithm 2-separate separates a set of n points P in the plane by a set H of $\lceil \frac{n}{2} \rceil$ lines.

Proof. After Step 1, P_1 and P_2 each contain at most $\lceil \frac{n}{2} \rceil$ points. Notice that any set of lines L' that separates both P_1 and P_2 together with ∂ separates P . We

know by Observation 4.3 that $H - \partial$ separates both of P_1 and P_2 . Also, by Observation 4.3, $|H| = (|P_2| - 1) + 1 = \lceil \frac{n}{2} \rceil$.

♦

Theorem 4.6. Algorithm 2-separate separates an n point set P in time $O(n^2)$ in the worst case and $O(n \log(n))$ in the expected case.

Proof. Clearly Step 1 is $O(n)$. Each 2-refine is really two 1-refines, one of P_1 and one of P_2 , that share the line used to refine the clusters. Also, by Theorem 4.4, it is clear that the choice of the line in 1-refine has no impact on the expected or worst case performance of 1-separate. Thus it follows that Algorithm 2-separate runs in time $O(n^2)$ in the worst case and $O(n \log(n))$ in the expected case.

4.4 Separating Points by Hyperplanes in E^d

We now consider separating points in E^d . We will first extend Algorithm 2-separate to work in higher dimensions. We will then consider some improvements to this extended algorithm to reduce the number of hyperplanes needed to separate points in E^d .

4.4.1 Extending Algorithm 2-Separate to Higher Dimensions

We now extend Algorithm 2-separate to work in arbitrary dimensions. The extension is straightforward and all the performance characteristics of Algorithm 2-separate are maintained. With the extended algorithm we show that $f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$.

We first observe that any set of $2k$ points in E^d $k \leq d$, grouped into k pairs $S = \{s_i \mid 1 \leq i \leq k\}$ in E^d , there exists a hyperplane H that simultaneously intersects the interior

of each of the k line segments determined by each pair of S . It is easy to show that such a hyperplane can be found in time $O(d^3)$, i.e. in constant time for any fixed d . An example, when $k = d$, is the hyperplane determined by the midpoints of these line segments or a small perturbation of it. See Figure 4.6 (b,c). This observation gives us a simple algorithm for carrying out a k -refine, $k \leq d$.

Algorithm k -refine

Input: A cluster set S_j of a point set P in E^d , and a subset $\{C_1 \dots C_k\}$ of k clusters of S_j , $1 \leq k \leq d$, $|C_i| \geq 2$, $1 \leq i \leq k$.

Output: 1) A cluster set S_{j+1} that results from a k -refine of C_i , $1 \leq i \leq k$, by a hyperplane h .

2) The hyperplane h .

- 1) For each C_i , $1 \leq i \leq k$, select any two points $p_{i1}, p_{i2} \in C_i$.
- 2) $h :=$ the hyperplane determined by the midpoints of $\text{seg}(p_{i1}, p_{i2})$, $1 \leq i \leq k$, and $d - k$ arbitrarily chosen points (or a small perturbation of this hyperplane).
- 3) For $i := 1$ to k do
 - 3.1) Partition C_i into two clusters C_{i1} and C_{i2} according to which side of h each point of C_i is on. /* A 1-refine of C_i */
- 4) Return($S_{j+1} := S_j - \{C_i, 1 \leq i \leq k\} \cup \{C_{ij}, 1 \leq i \leq k, 1 \leq j \leq 2\}, h$).

See Figure 4.6 (b,c).

If a set of hyperplanes H partitions a point set P in E^d into points sets of cardinality $\lfloor \frac{n}{t} \rfloor$ or $\lceil \frac{n}{t} \rceil$ then we say that H t -partitions P . The following result, to be proved later, will be used now.

Theorem 4.7. Any set P of n points in E^d can be $2^{\lceil \log(d) \rceil}$ partitioned using at most $\lceil \log(d) \rceil$ hyperplanes.

Once we construct d or more initial clusters (which we know we can do by Theorem 4.7), we can do simultaneous 1-refines of d subclusters, where each subcluster is a subset of a different one of the initial clusters, by using Algorithm d-refine. This leads us to our third point separation algorithm. (See Figure 4.6).

Algorithm d-Separate.

Input: A point set P in general position in E^d .

Output: A set of hyperplanes H that separates P .

- 1) $H_{\text{init}} := \{h_1 \dots h_{\lceil \log(d) \rceil}\}$ where $\{h_1 \dots h_{\lceil \log(d) \rceil}\}$ is a set of hyperplanes that $2^{\lceil \log(d) \rceil}$ partitions P into initial clusters C_i , $0 \leq i < 2^{\lceil \log(d) \rceil}$. We label the initial clusters such that $|C_i| \geq |C_{i+1}|$, $0 \leq i < 2^{\lceil \log(d) \rceil} - 1$. The initial clusters are put in a queue Q such that C_i precedes C_{i+1} , $0 \leq i < 2^{\lceil \log(d) \rceil} - 1$ in Q .
- 2) $S_0 := \{C_i, 0 \leq i < 2^{\lceil \log(d) \rceil}\}$, $H_0 := H_{\text{init}}$, $k := 1$.
- 3) While there is a crowded cluster D_j , $1 \leq j \leq d$ contained in each of the first d initial clusters in Q do
 - 3.1) find a hyperplane h_k that simultaneously splits D_j , $1 \leq j \leq d$ and then split them using h_k . This is accomplished using d-refine.
 - 3.2) $S_k = S_{k-1} - \{D_j, 1 \leq j \leq d\} \cup \{\text{the new clusters created by the d-refine}\}$
 - 3.3) Move the first d initial clusters in Q to the back of Q , maintaining their relative order.
 - 3.4) $H_k := H_{k-1} \cup h_k$, $k := k + 1$.

- 4) If there are m crowded clusters D_j , $1 \leq j \leq m$, remaining, then split each of them into two clusters using k -refine. $H_k := H_{k-1} \cup h_k$ where h_k is the hyperplane produced by the k -refine.
- 5) return H .

We call the clusters constructed by Step 1 of d -separate *initial clusters*. See Figure 4.6 (a).

Theorem 4.8. Algorithm d -separate separates a set of n points P in E^d by a set of hyperplanes H where $|H| = \left\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \right\rceil + \lceil \log(d) \rceil$.

Proof. After Step 1 any set of hyperplanes H' that separates C_i , $0 \leq i < 2^{\lceil \log(d) \rceil}$, together with H_{init} separates P . Let $C_{i,k}$, $0 \leq i < 2^{\lceil \log(d) \rceil}$, be the cluster set determined by the cluster C_i and the k -refines that have been applied to subclusters of C_i by the first k hyperplanes produced by d -separate after Step 1. Note that if the k th hyperplane does not split a subcluster of C_i then $C_{i,k} = C_{i,k-1}$. Let $R1(X)$ denote the number of k -refines necessary to separate cluster set X . By Observation 4.3 $R1(C_{i,0}) = |C_i| - 1$, $0 \leq i < 2^{\lceil \log(d) \rceil}$ and thus $R1(S_0) = \sum_{i=0}^{2^{\lceil \log(d) \rceil}-1} |C_i| - 1 = n - 2^{\lceil \log(d) \rceil}$.

Therefore S_0 can be separated using $\left\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \right\rceil$ hyperplanes if each hyperplane h_k , $1 \leq k \leq \left\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \right\rceil$, except possibly the last, is produced by a d -refine of d clusters of S_{k-1} . This is accomplished by d -separate since $|R1(C_{i,k}) - R1(C_{j,k})| \leq 1$, $0 \leq i, j < 2^{\lceil \log(d) \rceil}$ after Step 1 (when $k = 0$) and this equation remains invariant as hyperplanes are added by d -separate. Therefore the total number of hyperplanes produced by d -separate to separate P is $\left\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \right\rceil + \lceil \log(d) \rceil$.

♦

(a) the four initial cells determined by the two hyperplanes of step 1

(b) the first 4-split

(c) the second 4-split

(d) 20 points separated by 6 hyperplanes

The first three steps and the final set of hyperplanes produced by Algorithm 4-separate

Figure 4.6

Theorem 4.9. Not including Step 1 Algorithm d-Separate runs in time $O(n^2)$ in the worst case and $O(n \log(n))$ in the expected case.

Proof. By a slight generalization of Theorem 4.4 and by Theorem 4.6 it is clear that the cost of d-Separate is essentially the cost of doing a 1-separate of each C_i , $1 \leq i \leq 2^{\lceil \log(d) \rceil}$. In the worst case this cost is $O(d * \left(\frac{n}{d}\right)^2) = O(n^2)$. In the expected case it is $O\left(\frac{n}{d} * \log\left(\frac{n}{d}\right)\right) = O(n \log(n))$.

♦

4.4.2 Efficient k-Partitioning of n Points in E^d where $k \geq d$

We now show that it is possible to perform Step 1 of d-Separate. We also consider some extensions to Step 1 to reduce the number of hyperplanes required by d-separate and consider some variations to reduce the cost of Step 1.

We must point out that many of the "new" results that we present in this section were actually first presented in [Avis1] and [Avis2] and discovered independently by us. In particular, Theorems, Corollaries, and Open problems 4.11, 4.7, 4.14, 4.15, 4.16 and 4.18 are also wholly or partly presented in Avis's papers. Our presentation of these results, however, is quite different. For example the properties of the binary matrices used to prove theorem 6 of [Avis1] correspond to the properties of the balanced hamiltonian cycles for hypercubes that we discuss in this section, even though at first glance they appear quite different.

That Step 1 of d -separate is always possible is shown by the following theorem [Bors], [Will], [Schu].

Theorem 4.10. Let $P_1, \dots, P_d$ be d finite sets of points on E^d . There exists a hyperplane h that simultaneously bisects $P_1, \dots, P_d$.

We refer to any algorithm that finds a hyperplane that simultaneously bisects a collection of finite point sets in E^d as d -bisect. Consider a point set P in E^d and a set of hyperplanes $H = \{h_1, \dots, h_k\}$ that 2^k -partitions P . Let S_i , $0 \leq i \leq k$, be a family of cluster sets such that $S_0 = \{P\}$ and S_j is the cluster set determined by splitting each cluster of S_{j-1} with hyperplane h_j , $1 \leq j \leq k$. If h_j , $1 \leq j \leq k$, bisects each cluster of S_{j-1} then we say H is a k -bisector set of P . Clearly each cluster determined by any k -bisector set of P has cardinality $\lfloor \frac{n}{2^{\lceil \log(d) \rceil}} \rfloor$ or $\lceil \frac{n}{2^{\lceil \log(d) \rceil}} \rceil$. We are now ready to prove Theorem 4.7.

Proof of Theorem 4.7. Let P be any n point set in E^d . We need to show that there exists a set of $\lceil \log(d) \rceil$ hyperplanes that $2^{\lceil \log(d) \rceil}$ -partitions P . Clearly any $\lceil \log(d) \rceil$ -bisector set $H = \{h_1, \dots, h_{2^{\lceil \log(d) \rceil}}\}$ of P does this since each cluster of P determined by the cells of E^d , which are in turn determined by H , has cardinality $\lfloor \frac{n}{2^{\lceil \log(d) \rceil}} \rfloor$ or $\lceil \frac{n}{2^{\lceil \log(d) \rceil}} \rceil$. We need only show that H always exists. But by Theorem 4.9 and by the construction of H it is clear that each h_i , $0 \leq i \leq \lceil \log(d) \rceil$ exists and thus H exists.

♦

For E^2 there is a linear algorithm for d -bisect [Megi]. If we use this algorithm for d -bisect in step 1 of d -separate then the worst and expected run time behavior of d -separate is unchanged. For dimension three the best known algorithm for d -bisect is

$O\left(\frac{n^{3/2}\log^2(n)}{\log^*(n)}\right)$ [LMS]. Using this algorithm in step 1 of d -separate increases the expected run time behavior of d -separate but the worst case run time behavior is unaffected. We know of no algorithms for d -bisect for $d \geq 4$.

Step 1 of d -separate can be done in time $O(n)$ for arbitrary dimensions if we allow the use of more hyperplanes. Observe that an n point set P in E^d can be d -partitioned using $d-1$ hyperplanes all normal to the first coordinate axis. This can be done in $O(d * n)$ time since the k th element in an n element list can be found in $O(n)$ time.

We now explore the possibility of extending Theorem 4.7 so as to reduce the upper bound on $f(n,d)$ further. The upper bound on $f(n,d)$ can be reduced further for higher dimensions if Theorem 4.7 can be extended to allow 2^k -partitions of a point set in E^d where $2^k > \lceil \log(d) \rceil$. We thus pose the following open problem.

Open Problem 4.11. For each $d, d \geq 4$, determine the largest k such that for any n point set P in general position in E^d , $n \geq 2^k$, P can be 2^k -partitioned using k hyperplanes.

Let us denote such k by $bisect(d)$. For the above problem may be possible that $bisect(d) > d$ even though there does not always exist a hyperplane that can simultaneously bisect more than d point sets in E^d . Let H be a set of $bisect(d)$ hyperplanes that $2^{bisect(d)}$ -partitions P . Consider that the hyperplanes of H are applied to P in some order and let $h \in H$ be the last such hyperplane applied. Then h bisects $2^{bisect(d)-1}$ point sets. It may be that $bisect(d) > d$ for some values of d because the point sets bisected by h are not arbitrary but are determined by the other $bisect(d)-1$ hyperplanes of H . For example the following theorem [Bors],[Will],[Schu],[Edel] shows that $bisect(d) > d$ when $d = 3$.

Theorem 4.12. Any set of n points P in E^d , $d \geq 3$, can be 8-partitioned using three hyperplanes. One of the hyperplanes can be any bisector of P .

Once we know $\text{bisect}(d)$ for some dimension d we can use it in Step 1 of Algorithm d -separate. Step 1 then produces $2^{\text{bisect}(d)} - 1$ 1-refines. Therefore, in total, d -separate produces $\lceil \frac{n - 2^{\text{bisect}(d)}}{d} \rceil + \text{bisect}(d)$ hyperplanes that separate P . Thus $f(n, d) \leq \lceil \frac{n - 2^{\text{bisect}(d)}}{d} \rceil + \text{bisect}(d)$.

For example consider any 3-bisector set $H = \{h_1, h_2, h_3\}$ of an n point set P in E^3 , $n \geq 8$. Then h_3 does a 1-split of each of the four clusters of P determined by h_1, h_2 . Thus if we 8-partition P in Step 1 of 3-separate we will need to do four less 1-splits in the remaining steps of 3-separate than if we had 4-partitioned P . Since the remaining hyperplanes produced by 3-separate each produces at most three 1-splits it is clear that, for $n \geq 8$, 3-separate will produce $\lceil \frac{n - 2^3}{3} \rceil + 3$ hyperplanes. Simplifying gives us the following result.

Theorem 4.13. For $n \geq 8$, $f(n, 3) \leq \lceil \frac{n+1}{3} \rceil$.

The improvement we have made for dimension three is small. However each hyperplane added to a bisector set H of a point set P approximately doubles the number of 1-refines of P that are done by the hyperplanes of H . So it is clear that the improvement could be much larger in higher dimensions.

We can set bounds on the values of $\text{bisect}(d)$. We know already that $\text{bisect}(d) \geq \lceil \log(d) \rceil$. The following theorem determines an upper bound on $\text{bisect}(d)$.

Theorem 4.14. $\text{bisect}(d) \leq k$ where k is the largest value such that $kd > 2^k$.

Proof. Consider the moment curve $M_d(t) = (t, t^2, \dots, t^d)$ in E^d and let $\Psi_n(d) = \{M_d(1), \dots, M_d(n)\}$ on $M(t)$ be a set of n points on M_d . By Lemma 4.1, a single hyperplane can intersect M_d at most d times. Therefore any k hyperplanes H can intersect M_d at most kd times and thus split M_d into at most $kd+1$ sections. H therefore partitions $\Psi_n(d)$ into at most $kd+1$ clusters. Therefore k must be sufficiently small that $kd > 2^k$.

♦

Corollary 4.15. If $kd = 2^k$ for some k then $\text{bisect}(d) \leq k$.

Proof. In this case H still partitions M_d into 2^{k+1} sections. However the first section and the last section are on the same side of each hyperplane in H (since each hyperplane in H has intersected M_d d times between the first and last sections and d is even whenever $kd = 2^k$). Then these sections together correspond to at most one of the clusters of $\Psi_n(d)$ determined by H . Therefore H partitions $\Psi_n(d)$ into at most 2^k clusters and thus $\text{bisect}(d) \leq k$.

♦

The dimensions for which $kd = 2^k$ have a particularly interesting property which we show now. Let HC_r be the r -hypercube graph such that $V(HC_r) = \{v_0, \dots, v_{2^r-1}\}$ and for each vertex pair v_a and v_b , $a < b$ of $V(HC_r)$ there is an edge joining vertices v_a and v_b , if $a - b = 2^s$ for some s , $0 \leq s < r$. Let $E_t(HC_r)$, $0 \leq t < r$, be the set of edges of HC_r such that, for some vertex pair v_a and v_b , $a < b$ of $V(HC_r)$, $a - b = 2^t$. We observe that removing edges $E_t(HC_r)$ from HC_r breaks HC_r into two connected components each of which is a hypercube of dimension $r-1$. Thus HC_r determines r

pairs of hypercubes of dimension $r-1$. Clearly for each vertex v of $V(HC_r)$ the set of hypercubes of dimension $r-1$ that contain it is unique. It is easily shown that HC_r admits a hamiltonian cycle. A hamiltonian cycle through HC_r is said to be *balanced* if it crosses exactly r edges out of each set of $E_t(HC_r)$ edges, $0 \leq t < r$.

Theorem 4.16. Consider any dimension d such that $kd = 2^k$ for some k . Then $\text{bisect}(d) < k$ unless HC_k admits a balanced hamiltonian cycle.

Proof. Assume $\Psi_n(d)$ is partitioned into 2^k clusters $SC = \{C_1, C_2, \dots, C_{2^k}\}$ by some set of hyperplanes H . By Corollary 4.14 each section of M_d , determined by H , corresponds to a cluster of SC , except the first and last sections, which together correspond to a single cluster of SC . Let G be a graph in which each vertex v_i of G corresponds to cluster C_i of SC and in which there is an edge between vertices whose corresponding clusters are separated by a single hyperplane. Let λ be a path that starts at a point in the first section of M_d , determined by H , and follows M_d until the last section of M_d , determined by H , is reached. Path λ corresponds to a hamiltonian cycle, hc , through G since each section of M_d corresponds to a cluster which in turn corresponds a vertex of G and adjacent sections of M_d similarly correspond to adjacent vertices of G . We now show that G is a hypercube.

Define for each hyperplane $h_i \in H$ the halfspaces h_i^+ and h_i^- determined by h_i (we do not care which is which). Associate with each vertex v_i of G the 0-1 vector $X_i = \{x_1, \dots, x_k\}$ such that for x_j of X_i , $1 \leq j \leq k$, $x_j = 1$ if cluster C_i is in h_k^+ and $x_j = 0$ otherwise. Now, clearly the 0-1 vector for each vertex of G is unique. Furthermore, for any v_i, v_j of G , v_i and v_j are adjacent if and only if their 0-1 vectors differ in exactly one column. It is now clear that, after possible relabeling of the vertices, V is a hypercube.

Therefore if $kd = 2^k$ for some k and $\text{bisect}(d) = k$ then G admits a balanced cycle. Conversely, if $kd = 2^k$ for some k and G does not admit a balanced cycle then $\text{bisect}(d) < k$.

♦

For the values of d such that $kd = 2^k$ for some k and any n point set P in E^d can be 2^k -partitioned $f(n,d) \leq \lceil \frac{n-kd}{d} \rceil + k = \lceil \frac{n}{d} \rceil$. For these values of d $f(n,d) = \lceil \frac{n}{d} \rceil$. Since the bounds are then tight for $f(n,d)$, for these values of d , it is interesting to ask if we can increase the lower bound of $f(n,d)$ slightly and thus show that when $kd = 2^k$ for some k that $\text{bisect}(d) < k$. The method used to determine the lower bound for $f(n,d)$ is unlikely to help us, however, as we will show in the next theorem.

Let G be any curve in E^d such that any d distinct points on G determine a hyperplane h that intersects G exactly d times. Then G is called a *general position curve* on E^d .

Theorem 4.17. If $kd = 2^k$ for some k and HC_k admits a balanced hamiltonian cycle then for any n points P on any general position curve G on E^d there exists a set of k hyperplanes $H = \{h_1 \dots h_k\}$ that 2^k -partitions P .

Proof. Let $X = \{x_i, 0 \leq i \leq 2^k\}$, be a set of points on G such that the section of G between x_1 and x_{2^k} contains P and the section of G between any pair $(x_i, x_{i+1}), 0 \leq i \leq 2^k$, contains a set of $\lfloor \frac{n}{2^k} \rfloor$ or $\lceil \frac{n}{2^k} \rceil$ points C_i of P . Then $SC = \{C_1, C_2, \dots, C_{2^k}\}$ is a decomposition of P into 2^k subsets. (SC will eventually be the cluster set of P determined by H .) Then X partitions G into 2^k+1 sections such that each section, except the last, contains an element of SC . We associate the path λ from x_0 to x_{2^k} along G with a hamiltonian cycle, hc , such that x_i corresponds to the i th edge of hc

and C_i corresponds to the i th vertex of hc , $1 \leq i \leq 2^k$. (We can view λ as a cycle by considering x_0 and x_{2^k} to be the same vertex.)

We will now construct H . For each h_i , $1 \leq i \leq k$ we will determine the k points $p(h_i)$, of G that determine h_i . The following algorithm constructs H .

- 1) For $i := 1$ to k do $p(h_i) := \emptyset$.
- 2) For $i := 1$ to k do
 - 1.1) $e_i :=$ the i th edge of hc .
 - 1.2) $f := g$ such that $e_i \in E_g(HC_k)$.
 - 1.3) $p(h_f) := p(h_f) \cup x_i$.
- 3) end.

It is clear that this algorithm constructs H such that each h_i , $1 \leq i \leq k$, separates SC into two halves such that the point sets of one half correspond to the vertices of one of the two hypercubes of dimension $k-1$ of HC_k determined by $E_i(HC_k)$ and the other half corresponds to the vertices of the other hypercube. Therefore each C_i , $1 \leq i \leq 2^k$, is in a different one of the cells determined by H .

♦

If we attempt to improve the lower bounds of $f(n,d)$ by placing an n point set P on a curve which is not a general position curve then we have to deal with the possibility of hyperplanes that intersect the curve more than d times. But the lower bound for $f(n,d)$ (Theorem 4.2) was determined by choosing a curve explicitly so this could not happen. Thus a different approach will likely be needed to improve the lower bound on $f(n,d)$.

The above theorem leads to an open problem which is interesting in its own right.

Open Problem 4.18. For which dimensions d does a d dimensional hypercube admit a balanced hamiltonian cycle?

For HC_2 it is trivial to construct a balanced hamiltonian cycle. It is also clear that d must be even since any cycle must switch from one side to the other of any pair of hypercubes of dimension $d-1$ an even number of times. Figure 4.7 shows that we can construct a balanced hamiltonian cycle through HC_4 . With this hamiltonian cycle we can find a set of four hyperplanes H that partitions any given point set P on any general position curve G using the algorithm given in Theorem 4.17. See Figure 4.8.

The next two values of d for which $kd = 2^k$ for some k are 32 and 4096. The balanced hamiltonian cycle through HC_4 was easily found by a manual search. Determining whether or not there is a hamiltonian cycle in HC_{32} or HC_{4096} will clearly be considerably more difficult so that other approaches will be needed.

If $kd \neq 2^k$ for any k then we can still construct a path along a general position curve G that goes through a point set P on G in E^d that corresponds with a path p through HC_k . However p need not be a cycle and it may visit a node more than once. Nevertheless it may be possible to show that for some d that no such path through HC_d exists and thus reduce the upper bound on $\text{bisect}(d)$.

4.5 Separating Points Efficiently in E^2 and E^3

We show here that 2-separate can be improved to run in time $O(n \log(n))$ and that 3-separate can be improved to run in time $O\left(\frac{n^{3/2} \log^3(n)}{\log^*(n)}\right)$. We observe first that algorithm

Quicksort runs in time $O(n \log(n))$ if subalgorithm Partition selects the midpoint of each subcollection as the value to use to partition the subcollection. We will study improving Algorithm d-separate in an analogous way. We say that a hyperplane h bisects a cluster C containing n points if at most $\lceil \frac{n}{2} \rceil$ of the points of C are on each side of h . The change to 1-refine analogous to finding the midpoint in subalgorithm Partition then is to have it produce a hyperplane h that bisects the cluster being refined. For k -refine the analogous change is to have h simultaneously bisect each of the clusters being refined. We know from Theorem 4.10 that this is always possible.

We refer to any version of Algorithm k -refine which always produces a hyperplane h that bisects each cluster being refined k -bisect. If we simply replace Algorithm k -refine by k -bisect in Algorithm k -separate we call the Algorithm d -bisect-separate.

In E^2 linear algorithms for k -bisect are known [Megi]. If we use these algorithms then 2-bisect-separate runs in $O(n \log(n))$ time. This is because each point partakes in $O(\log(n))$ k -refines (k -bisects).

In E^3 there is an $O\left(\frac{n^{3/2} \log^2(n)}{\log^*(n)}\right)$ algorithm for doing a k -bisect [LMS]. Using this algorithm for k -bisect results in a time complexity of $O\left(\frac{n^{3/2} \log^3(n)}{\log^*(n)}\right)$ for 3-bisect-separate. For dimensions four and higher we know of no algorithms which do a k -bisect of a set of points, even though we know, by Theorem 4.10, that it is possible. In

[Cole1] and [Cole2] partitioning points in dimensions four and higher is studied. However his techniques require the use of considerably more hyperplanes than we know is required.

We conjecture that the problem of separating points in E^2 is $\Omega(n \log(n))$. However we have not been able to prove it. Proving it may be difficult since there does not appear to be any way to use the information gathered by separating a set of points in E^2 to solve a problem that is known to be $\Omega(n \log(n))$ such as sorting the points (say by x -coordinate).

4.6 Data Structures for Retrieval of Separating Hyperplanes and Lines

Now consider that d -separate has separated n points P by a set of hyperplanes H . We are then asked: given any two points $a, b \in P$, find a hyperplane of H that separates a, b . We now discuss data structures for storing the hyperplanes of H so that this query can be done efficiently. We consider first the case that the dimension is arbitrary. We show that, at least in theory, we can construct a data structure that allows a hyperplane that separates a, b to be found in time $O(\log(n))$. We then consider the planar case. Here we show that a line that separates a, b can be found in worst case time $O(\log(\log(n)))$ and expected time $O(1)$. In both cases we need a linear amount of space to store the hyperplanes.

4.6.1 A Data Structure for Retrieval of Separating Hyperplanes

There is a straightforward way to represent H so that a hyperplane that separates a, b can be found. We simply build a forest of binary trees with one binary tree for each initial cluster determined by Step 1 of Algorithm d -separate. Each node in each tree contains

one hyperplane, except for the leaf nodes, which contain no hyperplanes. Each edge of each tree represents the cluster determined by the initial cluster corresponding to the tree and by the hyperplanes on the path from the root to the edge. We note that each hyperplane, except those that determine the initial clusters, is duplicated in each tree. We call this structure a *hyperplane forest*. Figure 4.9 shows the hyperplane forest corresponding to Figure 4.6. A hyperplane forest is easily constructed during d -separate. Each of the d 1-refines of a cluster by d -refine corresponds to splitting a leaf of one of the d trees into an internal node and two new leaves.

If a, b are in different initial clusters then finding a separating hyperplane is straightforward. If a, b are in the same initial cluster then a hyperplane that separates them is found by searching the tree corresponding to the initial cluster that contains a, b . The search starts at the root and branches at each node according to which side of the hyperplane h (say) found at that node, a, b are on. If a, b are found on different sides of h , which must happen for some hyperplane during the search, then h separates a, b and the search terminates. We refer to the algorithm that performs this query *d-search*. Clearly d -search has run time $O(\text{the depth of the tree being searched})$. If the hyperplane forest is constructed by d -separate then d -search runs in time $O(n)$ in the worst case and $O(\log(n))$ in the expected case. If d -bisect-separate is used to construct the hyperplane forest then d -search runs in time $O(\log(n))$ since the no point partakes in more than $\lceil \log(n) \rceil$ bisections.

The binary forest of hyperplanes corresponding
to the set of hyperplanes in Figure 4.6

Figure 4.9

4.6.2 A Data Structure for Retrieval of Separating Lines

In the plane we can reduce our query cost to $O(\log(n))$ worst case and $O(1)$, expected case. To do this we will associate with each point $p_i \in P$ (which is a cluster) a convex

polygon or convex chain determined by a subset of lines of H . We will show that the sum, over all elements of P , of the sides of these polygons or chains is at most $4n$. We will also show that we can limit the maximum number of edges on any such polygon to $\lceil \log(n) \rceil$. This in turn leads to our claimed result.

In the planar case cells reduce to convex polygons, bounded or unbounded. Let H be a set of lines that separate a point set P . Observe that each cluster C has a cell associated with it; that determined by the subset of H which determined C ; We refer to that polygon as $CL(C)$. When a cluster C is refined by a line λ it is split into two clusters C_1, C_2 . $CL(C)$ is then split by λ into two new polygons $CL(C_1)$ and $CL(C_2)$. Clearly $|CL(C)| + 4 \geq |CL(C_1)| + |CL(C_2)|$. Since the number of 1-splits is $n-1$, the number of edges needed to construct $CL(p_i)$ for every separated point $p_i \in P$ (each of which is a cluster) is less than $4n$ and the amount of storage required to store these boundaries is $O(n)$. See Figure 4.10.

We can now find a separating line between any two points $a, b \in P$ very quickly. We simply determine the crossing support lines for $CL(a)$ and $CL(b)$. It is clear from the proof of Lemma 2.5 that at least one of the edges (of the two boundaries) that intersects the crossing support lines determines a separating line between the boundaries. See Figure 2.4. It is well known that the crossing support lines of $CL(a)$ and $CL(b)$ can be found in time $O(\log(k))$ where k is the size of the larger of $CL(a)$ and $CL(b)$. Thus a separating line between a, b can be found in time $O(\log(k))$.

The polygons corresponding to each cluster after the addition of separating lines. The points and lines are as in Figure 4.5

Figure 4.10

If 2-separate is used to separate P then the maximum number of edges on $CL(p)$ for any point $p \in P$ is $O(n)$ since p can partake in $O(n)$ 1-refines and with each 1-refine a new edge can be added to $CL(C)$ where C is the cluster containing p . Thus the cost of finding a separating line separating a, b , in worst case, is $O(\log(n))$. If, instead, 2-bisect-separate is used to separate P then p can partake in at most $\lceil \log(n) \rceil$ 1-refines and therefore the maximum number of edges on $CL(p)$ is $\lceil \log(n) \rceil$. The cost of finding a separating line separating a, b , in worst case, is then $O(\log(\log(n)))$. Regardless of which

separation algorithm is used the expected cost of finding a separating line is $O(1)$ since the average number of edges of over all polygons is less than four.

To construct the boundaries we maintain each boundary as a circular linked list of lines (each line can be thought of as an edge). Initially there are only two boundaries with each containing the initial line used to bisect P . For each cluster that is 1-refined we search its boundary for the edge intersection(s) with the line that is splitting the cluster, split the boundary into two new parts, and construct the new boundaries. In the worst case this cost is $O(k)$ where k is the number of edges on the boundary being split. If 2-separate is used then in the cost of constructing the boundaries for every point of P is worst case time $O(n^2)$. This can only occur if 2-separate actually runs in $O(n^2)$ time and the splits of the cluster boundaries are extremely uneven. In the expected case the boundaries are constructed in time $O(n)$. To see this we observe that in the expected case the maximum number of splits any point of P partakes in is $O(\log(n))$. The expected case analysis using 2-separate is then analogous to the worst case analysis for constructing the boundaries using 2-bisect-separate. But we show below that this cost is $O(n)$.

The worst case cost of constructing the boundaries improves considerably if 2-bisect-separate is used. For simplicity we assume that $|P| = n$ is a power of two. If not we can always add sufficient points to P until it is so. Let $G(n,k)$ be the set of clusters produced after each point p of P has participated in k 1-refines. Then $|G(n,k)| = 2^{k+1}$ since each 1-refine corresponds to a bisection of the point set containing p . Let $g(n,k)$ be the total number of edges in all the boundaries of the clusters of $G(n,k)$. Then $g(n,k) \leq g(n,k-1) + 4 * 2^k$ since at most four new edges are added to cluster boundaries when a cluster is split and there are $|G(n,k-1)| = 2^k$ clusters to split to construct $G(n,k)$ from $G(n,k-1)$. From this recurrence relation we determine that $g(n,k) < 2^{k+3}$. Let m be the total number of boundary edges to be processed during 2-bisect-separate. Then $m <$

$\sum_{i=0}^{\lceil \log(n) \rceil} g(n,i) = \sum_{i=0}^{\lceil \log(n) \rceil} 2^{i+3} < 16n$. Therefore the boundary for each point in P is

constructed in time $O(n)$.

5 Summary and Open Problems

In this thesis we surveyed the literature on the separability of convex sets in Euclidean vector spaces. We expanded the results where suitable to make them clearer. We also improved the results wherever we were able to do so.

Finally we added a new chapter on the separability of convex sets: the separability of sets of points in arbitrary dimension. This chapter represents the first significant results on separability dealing with dimensions higher than two. They are also almost the only results in this field where the upper bound and the lower bound on the number of lines or hyperplanes needed can be considered as close. For dimension two the bounds are tight. For higher dimensions the bounds are different only by a value logarithmic in the dimension.

In Chapter two we studied the general problem of separating arbitrary convex sets in the plane. We presented a result by [Feje] which proves that in the plane any set of n disjoint convex sets can be separated by a set of at most $3n-6$ lines. We also considered the problem of separating at least one element of a collection F of n convex sets from as many elements of F as possible. We presented a result by [Tver] which shows that for any collection of five convex sets that there is always one that can be separated by a line from at least two. We then presented a result in [CCUZ] in which it was shown that there is always one convex set of F that can be separated from at least $\lceil \frac{n}{12} \rceil$ elements of F .

In Chapter three we studied the problem of separating special classes of convex sets in the plane. In particular we studied the problems of separating circles in the plane and of separating line segments in the plane. This chapter is almost entirely based on results in [CCUZ] and the improvements we have made. They showed that one cannot always find two pairs of circles out of a collection of n circles such that some line separates each pair. They also showed that one cannot always separate one circle from $\lceil \frac{n}{2} \rceil$ but that for any set of n circles there is always a line that separates one circle from at least $\lceil \frac{n}{4} \rceil - 1$. We improved this result to $\lfloor \frac{n}{3} \rfloor - 2$.

The results on line segments were then surveyed. A configuration of n disjoint line segments T_n was presented such that no line exists that has two segments from T_n on each side of it. Then a configuration of $3n$ line segments J_n was presented such that no line can separate one element of J_n from a subcollection of J_n containing more than $n+1$ line segments. Thus there are collections of n line segments no one of which can be separated by a line from $\lceil \frac{n}{3} \rceil + 1$. This result represents not only the best upper bound we know of for separating one line segment from many but for separating one convex set from many.

A lower bound on separating one line segment from many was also determined. It was shown that for any collection S of n disjoint line segments that one can always find a line that separates one element of S from a subcollection of S containing at least $\lceil \frac{n-1}{4} \rceil + 1$ segments. If the elements of S can be extended to be semilines so that they remain disjoint then it is shown that there is an element of S that can be separated by a line from a subcollection of S which contains at least $\lceil \frac{n}{3} \rceil + 1$ elements. This bound is tight.

Finally in chapter four we presented the results on the separability of point sets in arbitrary dimension. Except where noted, the results of this chapter are entirely new.

Let $f(n,d)$ be the minimum number of hyperplanes always sufficient and occasionally necessary to separate n points in E^d (dimension d) where $n \geq d$. Then in E^2 (i.e. the plane) we proved that $f(n,2) = \lceil \frac{n}{2} \rceil$. For dimensions higher than two we proved that $\lceil \frac{n-1}{d} \rceil \leq f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d odd and $\lceil \frac{n}{d} \rceil \leq f(n,d) \leq \lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$ for d even. The upper bounds on $f(n,d)$ were determined by finding algorithms that generate a solution within these upper bounds. These algorithms are efficient, ranging in complexity from $O(n \log(n))$ to $O(n^2)$. The $O(n^2)$ algorithm is very easy to implement and very fast in practice, having an expected complexity of $O(n \log(n))$. We also discussed some possible ways of reducing the upper bound on $f(n,d)$.

We then developed a data structure for storing the hyperplanes that separate a set of points so that retrieving a hyperplane in the data structure that separates two given points can be found in worst case time $O(n)$ and expected case time $O(\log(n))$. We also showed that in E^2 we can construct an even more efficient data structure that allows us to retrieve a hyperplane that separates two given points in worst case time $O(\log(n))$ and expected case time $O(1)$.

We defined an interesting concept: that of a "General Position Curve". Let G be any curve in E^d such that any d distinct points on G determine a hyperplane h that intersects G exactly d times. Then G is a general position curve. We consider it an interesting idea worthy of mathematical study. We also posed an interesting problem involving hypercubes: that of finding "balanced" hamiltonian cycles through hypercubes.

Table 5.1 summarizes the current best lower and upper bounds we know of for each of the problems studied. Not included are the results that 1) we cannot always separate two elements from two (for circles and for line segments) and that 2) out of five convex sets we can always separate one from two.

Class	Dimension	Number of objects separated	Number of objects separated from (# lines to separate all from all)	
			Lower bound	Upper bound
Convex Sets	2	1	$\lceil \frac{n}{12} \rceil$	$\lceil \frac{n}{3} \rceil + 1$
Convex Sets	2	all from all	$(3n-6)$	$(3n-6)$
Circles	2	1	$\lfloor \frac{n}{3} \rfloor - 2$	$\lceil \frac{n}{2} \rceil$
Congruent Circles	2	$\frac{n}{2} - c \sqrt{n} \sqrt{\log(n)}$	$\frac{n}{2} - c \sqrt{n} \sqrt{\log(n)}$	
Line Segments	2	1	$\lceil \frac{n-1}{4} \rceil + 1$	$\lceil \frac{n}{3} \rceil + 1$
Semilines	2	1	$\lceil \frac{n}{3} \rceil + 1$	$\lceil \frac{n}{3} \rceil + 1$
Points	2	all from all	$\lceil \frac{n}{2} \rceil$	$\lceil \frac{n}{2} \rceil$
Points	3	all from all	$\lceil \frac{n-1}{3} \rceil$	$\lceil \frac{n+1}{3} \rceil$
Points	4	all from all	$\lceil \frac{n}{4} \rceil$	$\lceil \frac{n}{4} \rceil + 1$
Points	d even	all from all	$\lceil \frac{n}{d} \rceil$	$\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$
Points	d odd	all from all	$\lceil \frac{n-1}{d} \rceil$	$\lceil \frac{n-2^{\lceil \log(d) \rceil}}{d} \rceil + \lceil \log(d) \rceil$

Summary of results of separating convex sets

Table 5.1

It is worth noting that the separation of points in E^d has applications in several geometric searching problems; see for example [Edel]. However, the results we presented in Chapter four are of no help in solving these problems.

5.1 Some Open Problems

We now discuss a number of interesting open problems. Of course each gap between columns four and five of table 5.1 implies the open problem of closing or at least reducing the gap. We think that the upper bound on separating one convex set from many is far too large. This upper bound was constructed by counting edges in a graph in such a way that it "feels" that we have missed many edges. We tried and so far failed to increase the number of edges counted so that the upper bound would be reduced to about $\lceil \frac{n}{6} \rceil$. Nevertheless this stronger result may still be achievable by this approach and it might even be able to reduce the bound even further.

The case of congruent circles is already a very strong result so that there is very little room for improvement. However the constant c is completely unspecified so that determining or at least bounding the value of c is an interesting (and probably difficult) problem.

The bounds on separating points in E^3 is almost tight and it would be nice to make them tight. In E^4 the obvious open problem is "can we separate n points into sixteen clusters of essentially the same cardinality?".

For points sets in higher dimensions there is not a lot of room for improvement. Nevertheless the upper bounds on separating points in arbitrary dimensions do not "feel" tight and it should be possible to reduce them at least marginally. We expect that the lower bounds for separating points in arbitrary dimension are probably close to the correct values.

We have seen for convex sets in E^d that we cannot, in general, find a hyperplane that separates one convex set from two other sets. However this might change if we use restricted classes of convex sets. We have seen already that for points the situation is completely different.

It might be possible to extend the proof of Theorem 3.3. to higher dimensions to prove that one hypersphere can be separated from $O(\binom{n}{d})$ hyperspheres. However this proof is dependant upon a large number of lemmas and it is not at all clear that they can (all) be generalized, in some form, in higher dimensions.

Consider, for example, Lemma 3.5. It says that if two circles have their centers in the same quadrants, intersect both axes, and neither circle contains the origin then the circle further away from the origin has a radius at least four times as large as the radius of the other circle. If we pose the corresponding problem for hyperspheres then we find that the ratio between the spheres is dependent upon the dimension and in fact, for dimensions greater than five, the ratio can be proved to be less than one! This difficulty can be dealt with but this lemma was the easiest of the many lemmas needed to prove Theorem 3.3. Even if we consider only the case of dimension three then generalizing the remaining lemmas is a difficult, though entirely geometrical, problem.

An interesting problem relates to that of general position curves described in Chapter four.

Open Problem 5.1. Let $F(d)$ be the maximal number of points in E^d such that for any collection of $F(d)$ points (not aligned) there always exists a general position curve that contains them. Determine $F(d)$.

For example, we can prove that $F(2) = 3$.

All of the results in this thesis find a solution for the worst case scenario. We could, however, also phrase the problem in terms of finding a solution as close as possible to the optimal solution for a given set of convex sets. Consider for example the case of separating points, say in E^2 . We could ask to find the minimal number of lines needed to separate a given set of points. We expect that this problem is NP-Complete. We might then try find a set of lines L such that $|L|$ is related to the size of the optimal solution, say within a fixed fraction of it.

References

- [AKP] N. Alon, M. Katchalski, and W.R. Pulleyblank. "Cutting disjoint disks by straight lines", *Discrete Comput. Geom.* **4** (1989), 239-243.
- [Avis1] Avis, D. "On the partitionability of point sets in space". Proceedings of the 1st A.C.M. Symposium on Computational Geometry. 1985, 116-120.
- [Avis2] Avis, D. "Non-Partitionable Point Sets" *Information Processing Letters* **19** (1984) 125-129
- [Bors] Borsuk, K. "Drei Sätze über die n -dimensionale euklidische Sphäre". *Fund. Math.* **20** (1933).
- [CCUZ] Rivera-Campo, E., Czyzowicz, J., Urrutia, J., and Zaks, J. "Separating Convex Sets On the Plane". *Discrete Comput. Geom.*, To appear.
- [Cole1] Cole, R. "Partitioning point sets in 4 dimensions". Proceedings 12th international Colloquium on Automation, Languages, and Programming. 1985, 111-119, *Lecture Notes in Computer Science* **194**, Springer-Verlag, Berlin
- [Cole2] Cole, R. "Partitioning point sets in arbitrary dimensions". *Theoretical Computer Science.* **49** (1987)
- [Edel] Edelsbrunner, H. *Algorithms in Combinatorial Geometry* Springer-Verlag, Berlin 1987.
- [EdWe] Edelsbrunner, H. and Welzl, E. "Halfplanar range search in linear space and $O(n^{0.695})$ query time", *Information processing letters* **23** (1986) 289-293.

- [ERS] Edelsbrunner, H., Robin A.D., and Shen, X. "Covering convex sets with non-overlapping polygons", Report No. UIUCDCS-R-87-1364. Aug. 1987.
- [Feje] Fejes Toth, L. "Illumination of convex disks." *Acta Mathematica Academiae Scientiarum Hungaricae* **29** (1977) 355-360.
- [Grün] Grünbaum, B. *Convex Polytopes*. John Wiley and Sons, London, 1967.
- [LMS] Lo, C., Matousek, J., and Steiger, W. "Ham-Sandwich Cuts in two and three dimensions". Technical report number ??? Department of Computer Science Rutgers University. August 1991.
- [Megi] Megiddo, N. "Partitioning with two lines in the plane". *Journal of Algorithms* **3** (1985) 430-433.
- [Schu] Schurle, A.W. *Topics in topology*. North-Holland, New York, 1979
- [Tver] Tverberg, H. "A separation property of plane convex sets", *Math. Scand.* **45** (1979), 255-260.
- [UrZa] Urrutia, J. and Zaks, J. "Illuminating convex sets", Technical Report TR-89-31. Department of Computer Science, University of Ottawa.
- [Will] Williard, D.E. Polygon retrieval. *SIAM Journal of Computing*. **11** (1982)

END

30.12.92

FIN