Discrete Comput Geom 9:257–266 (1993)

Discrete & Computational
Geometry

An Equipartition of Planar Sets*

Leonard J. Schulman

Department of Mathematics, Massachusetts Institute of Technology,
Cambridge, MA 02139, USA

Abstract. We describe the "cobweb" partition scheme and show that it can split any planar set into eight regions of equal area.

1. Introduction

It is commonly observed that any (measurable) set in the plane can be cut by a pair of lines into four parts of equal area. Courant and Robbins [4] showed a stronger theorem: that the pair of lines can be taken to be perpendicular. We rephrase this to say that some pair of perpendicular lines defines an *equipartition* of the set. The stronger result harnesses the intermediate value (Bolzano's) theorem to turn to advantage a degree of freedom which goes to waste in the proof of the weaker version (e.g., one line is often taken horizontal).

Buck and Buck [3] used the same degree of freedom to show that a convex planar set can be partitioned by three concurrent lines into six parts of equal area. As with the perpendicular line pairs, fundamentally the method of proof is to consider a continuous family of partitions, use a real-valued function to measure how close each partition is to being an equipartition, and use the intermediate value theorem to argue that for some member of the family this function is zero. Thus that member is an equipartition. Edelsbrunner and Huber [5] proved a discrete analog of this theorem; an analog for measurable sets (more generally, "density functions" as defined below) can also be observed.

We refer to various ways of cutting up the space—such as by pairs of lines, pairs of perpendicular lines, or three concurrent lines—as partition schemes. Loosely speaking this is a set of partitions of space, parametrized by some manifold. (For instance, partitions by line pairs are parametrized by $\mathbb{R}^2 \times (S^1)^2$).

*This research was supported by an ONR graduate fellowship.

Recent interest in equipartitions was stimulated by Willard's [10] recognition of their usefulness in the solution of "range search" problems. Willard's lead has been extended in many ways: the reader is directed to [12] for a discussion, and to [6] for some recent results. Equipartitions have also arisen in the study of "crossing families" [2].

Several investigations (see [12]) have been examined partition schemes for dimensions higher than two. For instance F. Yao *et al.* [13] showed that three planes can equipartition any set in three dimensions; while A. Yao and F. Yao [11] described certain partition schemes for arbitrary dimensions and showed that they equipartition every set. Broadly speaking these papers adopt the following approach (analogous to that described above, but using the Borsuk–Ulam theorem, a generalization of the intermediate value theorem) to establish their existence proofs: parametrize a set of constructions by the sphere S^k, and map the sphere onto $\mathbb{R}^k$ in such a manner that any construction with the desired property (e.g., equipartition) will be carried to the origin. Then show that such a construction exists by observing that the map is continuous and antipodal; and that therefore, by the Borsuk–Ulam theorem, the range of the map includes the origin.

As an example of the application of the Borsuk–Ulam theorem in this context, use the circle to parametrize (by orientation) the set of perpendicular line pairs, each of which evenly splits a given set in the plane. Now prove Courant and Robbin's theorem by considering the difference in the areas covered by two adjacent quadrants.

We describe the "cobweb" partition scheme for the plane. It uses two lines and four line segments to cut the plane into eight regions (Fig. 1). We show that it equipartitions any planar set.

As in previous investigations, we adopt the strategy of continuous parametrization. We rely on topological arguments twice. In one of these we merely apply the intermediate value theorem much as described above. In the other we show that a map between two tori is a homotopy equivalence by examining the induced map on their fundamental groups.

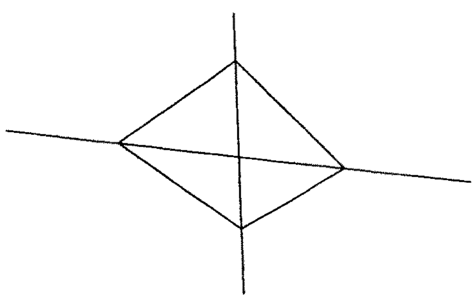

Fig. 1. A cobweb.

2. Preliminaries

Definitions

We first define some topological terms. (Useful references are [7] and [9]). The circle is denoted by S^1 and the real line by $\mathbb{R}$. Let X be a topological space. We say that X is *path connected* if for every pair of points $x, y \in X$ there is a path from x to y: i.e., a continuous map g: $[0, 1] \to X$ such that $g(0) = x$ and $g(1) = y$.

If f: $X \to \mathbb{R}$ is a nonnegative function, then the *support* of f is that subset of X on which f is positive. If this subset is compact, then we say that f is *compactly supported*. (We are concerned with the subsets of a Euclidean space $X = \mathbb{R}^k$. In this case a set is compact precisely if it is closed and bounded.)

Now we describe the class of functions we work with. Let μ be the Lebesgue measure (i.e., area function) in the plane. (References to measurability in this paper are necessary for technical reasons but do not otherwise feature in our arguments. They may be clarified in an analysis text such as [1]).

Definition. A function f: $\mathbb{R}^2 \to \mathbb{R}$ is a *density function* if it is nonnegative, measurable, compactly supported, and if there exists some real K such that $\int_B f < K\mu(B)$ for all measurable sets B.

Observe that the integral of a density function over any measurable set is finite. Let $\int_{\mathbb{R}^2} f = M$.

Definition. A partition of the plane into a regions is an *equipartition* of f if the integral of f over each region is M/a.

In order to partition a measurable subset of the plane (of finite measure, i.e., area) we represent the set by its characteristic function, which is 1 on the set and 0 otherwise.

Density Functions and Finite Sets

Before we show why any density function can be equipartitioned with a cobweb, let us note that this fact implies the same for a finite set of points in general position. Let S be such a set, and suppose that 8 divides $|S|$.

Thicken each point of S to a very small disk. Then define f_S by letting it be constant of integral 1 within each disk, and 0 outside the disks. Thus f_S is a density function, and $M = |S|$. Find an equipartition of f_S. This will be an equipartition of S unless perhaps if some of the disks intersect the lines and segments defining the partition. In this case the partition can be perturbed minutely to one in which every disk lies entirely in some region, without affecting the weights of the regions.

Thus the theorem on density functions (Section 3) implies:

Corollary. *Every finite set of points in general position in the plane is equipartitioned by some cobweb.*

Torus Maps and Rotating Pairs of Splitters

We have already alluded to the fact that for any density function f there exist a line l_1 of arbitrary orientation, and another line l_2, splitting f into four quadrants (not necessarily right-angled) of equal weight. Here is a typical argument: choose the orientation of l_1; displace it until it splits f; now start with an l_2 splitting one half-plane, and maintain this property while rotating l_2 until it also splits the other half-plane.

The neglected degree of freedom here is plainly evident in the orientation of l_1. We make essential use of this freedom to rotate l_1 and l_2 through an entire circuit, always splitting f into four equal parts, before returning to the starting position. The fact that this can be done in a continuous fashion is a key ingredient of the argument in the next section; we prove it in the present section. First, however, we provide an informal argument which the reader may find sufficient.

Let us suppose that f is nonzero only on some finite disk in the plane; and furthermore that within that disk it is continuous, and bounded below by some $\varepsilon > 0$.

Examine any specific l_1, l_2 which split f into four equal parts. It may be seen that on each of l_1, l_2 there is some point about which they can be infinitesimally rotated without imbalancing their split of the weight of f (i.e., if the difference in the weights of the half-planes is expressed as a function of the angle, then the derivative of this function is zero). It may also be seen that each of these rotations has a nontrivial effect on the balance of weights of the quadrants. (That is, if the difference between the weights of any pair of adjacent quadrants is expressed as a function of the angle of l_1, then the derivative of this function is nonzero. Similarly for l_2.) It therefore follows from the implicit function theorem [8, Section 9] that the imbalancing effect on the quadrants, of an infinitesimal rotation of l_1, can be compensated for by an infinitesimal rotation of l_2. This implicitly defined dependence of l_2 on l_1 is continuously differentiable, as l_1 varies over some small domain. (An open set in its parametrization space $\mathbb{R} \times S^1$.)

The local functions (for l_2 in terms of l_1) constructed in this way can be patched together to give a global function: so that, as l_1 is rotated through a full cycle, the l_1, l_2 configuration undergoes a continuous rotation, always maintaining an equipartition of f. At the end of the cycle, l_1 and l_2 return to their original positions.

The remainder of this section is devoted to a full proof, without unnecessary assumptions on f, that l_1, l_2 can be continuously rotated while maintaining an equipartition of f. Some homotopy-theoretic terms are needed only in this section, and can be found in [9].

Surround the convex hull of the support of f by a large circle O. Each directed line λ intersecting that convex hull meets O in two points, λ^+ and λ^-. Let $\mathscr{L}$ be the set of directed lines intersecting O. We topologize $\mathscr{L}$ with the following metric (where $\lambda_1, \lambda_2 \in \mathscr{L}$):

$$d(\lambda_1, \lambda_2) = \max_{v = \lambda_1^+, \lambda_1^-, \lambda_2^+, \lambda_2^-} (\text{distance from } v \text{ to the other line}).$$

We further restrict ourselves to the subspace $\mathscr{L}_*$ of $\mathscr{L}$, consisting of those lines which split f into halves of equal weight. Call these lines splitters. Observe that at every orientation, $\mathscr{L}_*$ restricts to one line or to a family of parallel lines parametrized by a closed interval. We claim that the set of all "rightmost" splitters (each representing such a family of parallel, commonly oriented splitters) is a loop (continuous map from S^1) in $\mathscr{L}$; as is the set of all "leftmost" splitters. We show this for the "rightmost" splitters, $\mathscr{L}_{**}$, as follows. Let m be the rightmost member of $\mathscr{L}_*$ at the orientation β. If there is a discontinuity at β in the map $S^1 \to \mathscr{L}_{**}$ sending an orientation to the rightmost splitter at that orientation, then there is an $\varepsilon > 0$ and a sequence of angles $\{\beta_i\}_1^\infty$ converging to β such that (where m_i is the rightmost splitter at angle β_i) $d(m_i, m) > \varepsilon$ for all i. For β_i very close to β this implies that the intersection of m_i and m must lie outside O. Then m_i is strictly to one side of m within O, and since both m and m_i are splitters, the band between them within O must have measure 0. However, then they cannot both be rightmost—for one of them can be displaced rightward into the band.

Proposition. *There are continuous maps, $l_1, l_2 \colon S^1 \to \mathscr{L}_{**}$ such that, for all $\alpha \in S^1$, the quadrants defined by $l_1(\alpha)$ and $l_2(\alpha)$ all have equal weight; and such that, as α traces out one circuit of S^1, the orientations of $l_1(\alpha)$ and $l_2(\alpha)$ each make exactly one net rotation.*

Proof. If λ_1 and λ_2 are intersecting directed lines, let $\lambda_{1,2}$ be their point of intersection, and let λ_1^+, λ_1^-, λ_2^+ and λ_2^- (as before) be their intersections with O. Let $\mathscr{T}$ be a torus.

We define a map $H\colon \mathscr{L}_{**} \times \mathscr{L}_{**} \to \mathscr{T}$, $H(\lambda_1, \lambda_2) = (\alpha, \eta)$ as follows. α is the orientation of λ_1. If λ_1 and λ_2 intersect, then η is $2\pi/M$ times the weight of the wedge-shaped region (quadrant) swept clockwise about $\lambda_{1,2}$ from λ_1^+ to λ_2^+. Otherwise if λ_1 and λ_2 have the same orientation, set $\eta = 0$, while if they have opposing orientations, set $\eta = \pi$. Since f is a density function, H is continuous.

The heart of our proof lies in showing that H is a homotopy equivalence. In order to see this we first examine the images of generators for the fundamental group $\pi_1(\mathscr{L}_{**} \times \mathscr{L}_{**})$. The fundamental group of a torus is isomorphic to Z^2. Let y be the path in $\mathscr{L}_{**} \times \mathscr{L}_{**}$ which, for some fixed λ_1, rotates λ_2 once counterclockwise. Then $H(y)$ is a path which leaves α fixed and makes one counterclockwise circuit in η. Let z be the path in $\mathscr{L}_{**} \times \mathscr{L}_{**}$ which rotates both λ_1 and λ_2 through one full counterclockwise circuit, while keeping them at some fixed angular separation. Then $H(z)$ is a path which makes one counterclockwise circuit in α, while making no net change in η. Hence $H(y)$ and $H(z)$ generate $\pi_1(\mathscr{T})$. We conclude that H induces an isomorphism of fundamental groups.

The torus is a triangulable space and therefore can be given the structure of a CW-complex. It is a theorem of J. H. C. Whitehead [9, Section 6] that if a continuous map among CW-complexes induces isomorphisms of all their homotopy groups, that map is a homotopy equivalence. Since the higher homotopy groups of the torus vanish, we find that H possesses a homotopy inverse $J: \mathscr{T} \to \mathscr{L}_{**} \times \mathscr{L}_{**}$.

Thus $J(H(y))$ and $J(H(z))$ are homotopic to y and z, and generate $\pi_1(\mathscr{L}_{**} \times \mathscr{L}_{**})$. Let u be a path in $\mathscr{T}$ which makes one counterclockwise circuit in α while keeping η at the fixed value $\pi/4$. u is homotopic to $H(z)$, therefore $J(u)$ is homotopic to $J(H(z))$ and, thus, to z. Hence $J(u)$ is our desired rotating pair of splitters l_1, l_2 which split f into equal-weight quadrants. □

3. Equipartition by a Cobweb

If a pair of splitters of a density function f should happen to separate the plane into quadrants of equal weight, then we refer to the splitters as a pair of *quartering axes* for f.

A *bisector* of a quadrant Q is any ray or line segment in Q, whose endpoint(s) lie on the half-axes bounding Q, and which separates Q into two regions of equal weight. Observe that any two bisectors of Q intersect unless the region between them is of weight 0.

A *diamond* is a set of four bisectors, one in each quadrant, such that endpoints of adjoining bisectors meet. We refer to the structure of a pair of quartering axes, with four bisectors in the shape of a diamond, as a cobweb (see Fig. 1).

Theorem. *Every density function in the plane is equipartitioned by some cobweb.*

Proof. We define a *pseudodiamond* as a sequence of five connected bisectors wrapping around the origin: the first and last are in the same quadrant, and the head of each touches the tail of the next. The first and last may have their free ends on the axes or at infinity. Let us denote the six points defining these bisectors as $v_0, \ldots, v_5$, with v_0 and v_5 possibly infinite. Denote the intersection of the axes by w.

Given the points $v_1, \ldots, v_4$ of a pseudodiamond, there may be some choice in v_0 and v_5 due to the existence of regions in which the density function, f, has weight 0. If we arbitrarily choose v_0 and v_5 to be as far inward (close to w) as possible, then v_4v_5 and v_1v_0 must intersect, and we find that there are just two types of pseudodiamonds (Fig. 2):

- *Inner pseudodiamonds*, where the free ends are finite, and hit their respective axes in between w and the other pseudodiamond point. (Thus v_0 is between w and v_4, and v_5 is between w and v_1.)
- *Outer pseudodiamonds*, where the free ends are either infinite or hit their respective axes beyond the other pseudodiamond point. (Thus v_0 is beyond v_4, and v_5 is beyond v_1.)

There is one boundary case contained in both of these types: when $v_0 = v_4$ and $v_1 = v_5$ we have a diamond.

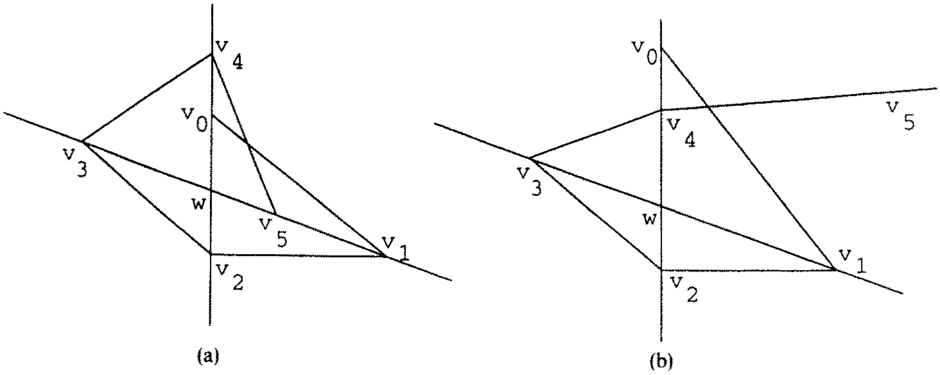

Fig. 2. (a) Inner and (b) outer pseudodiamond.

We refer to the segment v_2v_3 as the *base* of the pseudodiamond. Given a pair of quartering axes and a bisector of one of the quadrants, it is not always possible to construct a pseudodiamond with that bisector as its base. In order to understand when this is possible, we introduce the eight "standard" bisectors, those parallel to the quartering axes. In case there is any freedom in their placement, locate them as close to w as possible. Call their intersections with the axes, *notches*. We now show that the condition that a pseudodiamond exist with a given base v_2v_3 is that (see Fig. 3):

(I) v_2 lies strictly beyond the *adjoining* quadrant's notch on v_2's half-axis; and
(II) v_3 lies strictly beyond the *adjoining* quadrant's notch on v_3's half-axis.

In view of the fact that f is compactly supported, (I) ensures that v_1 can be chosen finite, and (II) does the same for v_4. This is all that is needed in order to guarantee that v_2v_3 is the base of a pseudodiamond. Conversely, if the condition is violated, v_1 and v_4 will not both be finite, and a pseudodiamond with base v_2v_3 will not exist.

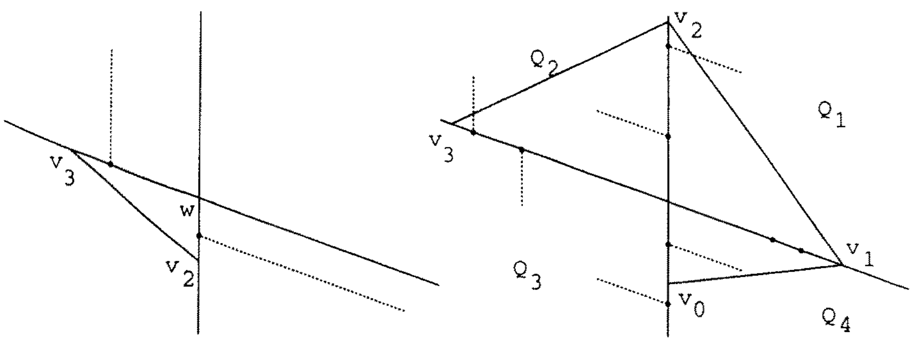

Fig. 3. Existence condition; Case 1.

We now establish:

Lemma. *For some choice of quartering axes, there exists an outer pseudodiamond with at least one finite endpoint.*

Proof. Among each pair of adjacent quadrants Q_i and Q_j, set $Q_i \to Q_j$ if the notch for Q_i on their common half-axis is further from w than is the notch for Q_j.

Two cases arise:

Case 1: *There exist consecutive quadrants* Q_1, Q_2, Q_3 *such that* $Q_1 \to Q_2 \to Q_3$. Pick a point v_1 on the half-axis between Q_4 and Q_1, beyond both of their notches on that half-axis (see Fig. 3). A bisector v_1v_0 can be extended through Q_4 to a finite point v_0. Similarly, a bisector v_1v_2 can be extended through Q_1 to a finite point v_2. v_2 must be at least as far out as the notch of Q_1 on that half-axis. Since $Q_1 \to Q_2$, v_2 is at least as far out as the relevant notch of Q_2. Therefore a bisector v_2v_3 may be extended through Q_2 to a finite v_3. v_3 must be at least as far out as the relevant notch of Q_2; and, since $Q_2 \to Q_3$, v_3 is at least as far out as the relevant notch of Q_3. Therefore a bisector v_3v_4 may be extended through Q_3 to a finite v_4. Finally, a bisector v_4v_5 may be extended through Q_4 to a finite or infinite v_5.

We have constructed a pseudodiamond with at least one finite endpoint, v_0. If the pseudodiamond is outer, we are done. Otherwise v_0 is between v_4 and w. A bisector extended from v_0 through Q_3 can be chosen to arrive at point v_{-1} outside v_3. $v_{-1}, \ldots, v_4$ is an outer pseudodiamond, with finite endpoint v_4.

Case 2: *A pair of opposing quadrants arrow the two others.* Rotate the axes. By the time one axis reaches the other's present position, the arrows will all have reversed. At the time of the first change we are in case 1. □

We prove the theorem by starting from an outer pseudodiamond with finite v_0, as provided by the lemma (see Fig. 2(b)). Rotate the axes in the direction from v_1 toward v_0. We now make several observations:

1. For any quartering axes and selected quadrant: (a) The range of angles available for the base of a pseudodiamond in that quadrant is parametrized by an open (possibly empty) interval. (b) The space of all allowable bases is path-connected.
   Proof: (a) We showed earlier that the condition (I, II) for the existence of a pseudodiamond was necessary and sufficient. The positions of the two notches discussed there, limit the angles available to the base of a pseudodiamond. (b) To transform one base into another, rotate it until a pair of endpoints identify; then the region remaining between the two bases is of weight 0 and can be closed continuously.
2. The bounds of the parametrizing interval vary continuously.
   Proof: The positions of the notches are continuous functions of the positions of the quartering axes, which in turn, as we demonstrated in the proposition earlier, can be rotated continuously.

3. The entire configuration of a pseudodiamond is a continuous function (unique up to some possible choice due to weightless regions) of the quartering axes and the base segment.
   Proof: v_0 can be chosen continuously in the quartering axes and v_1; v_1 can be chosen continuously in the quartering axes and v_2. Similar reasoning applies to v_5 and v_4.

These observations enable us to rotate the pseudodiamond continuously along with the quartering axes—always keeping the angle of the base away from the bounds of the parametrizing interval, as long as that remains nonempty.

Two possibilities arise. The first is that as the axes rotate, the parametrizing interval is always nonempty. Then the axis carrying v_1 reaches the position formerly occupied by the axis carrying v_0, with a continuous choice of pseudodiamond throughout. Now shift the base until it is in the exact position formerly occupied by the segment v_2v_1. That position is suitable for the base of a pseudodiamond because the original v_3 and v_0 were finite (Fig. 2(b)). Since the space of bases is path-connected, this shift can be accomplished continuously. Also (should this be left to choice due to weightless regions), shift v_1 until it is in the former position of v_0, v_4 of v_3, and v_5 of v_4.

Now recall that we started with an outer pseudodiamond: v_0 was outside of v_4. This means that, presently, v_1 is outside of v_5—indicating that we have arrived at an inner pseudodiamond. Having continuously moved from an outer to an inner pseudodiamond, somewhere on the way we must have encountered a diamond.

The second possibility is that at some moment during the rotation, the parametrizing interval vanishes. Then by stopping arbitrarily shortly before that moment, we can find a pseudodiamond with v_2 and v_3 each arbitrarily close to the notch into the adjoining quadrant. In that case v_1 and v_4 can each be taken arbitrarily far from w. Then since f is compactly supported, a bisector extended from v_1 must reach a point v_0 between w and v_4; and a bisector extended from v_4 must reach a point v_5 between w and v_1. Hence this pseudodiamond, which we have arrived at continuously, is inner. On the way we must have passed through a diamond. □

Acknowledgements

I thank Wayne Goddard for extensive discussions and criticisms; and Dan Kleitman, Mike Sipser and David Yavin for helpful discussions.

References

1. M. Adams and V. Guillemin. *Measure Theory and Probability*, Wadsworth, Monterey, CA, 1986.
2. B. Aronov, P. Erdös, W. Goddard, D. J. Kleitman, M. Klugerman, J. Pach, and L. J. Schulman. Crossing families. In *Proc. Seventh ACM Symp. on Computational Geometry*, pp. 351–356, 1991.
3. R. C. Buck and E. F. Buck. Equipartition of convex sets. *Math. Mag.*, 22:195–198, 1987.
4. R. Courant and H. Robbins. *What is Mathematics*, pp. 317–319. Oxford University Press, Oxford, 1941.

5. H. Edelsbrunner and F. Huber. Dissecting Sets of Points in Two and Three Dimensions. Technical Report F138, Technische Universitat Graz, 1984.
6. J. Matoušek. Efficient partition trees. In *Proc. Seventh ACM Symp. on Computational Geometry*, pp. 1–9, 1991.
7. J. R. Munkres. *Topology: A First Course*. Prentice-Hall, Englewood Cliffs, NJ, 1975.
8. W. Rudin. *Principles of Mathematical Analysis*, 3rd edn. McGraw-Hill, New York, 1976.
9. R. M. Switzer. *Algebraic Topology—Homotopy and Homology*. Springer-Verlag, Berlin, 1975.
10. D. E. Willard. Polygon retrieval. *SIAM J. Comput.*, 11(1):149–165, February 1982.
11. A. C. Yao and F. F. Yao. A general approach to *d*-dimensional geometric queries. In *Proc. 17th Symp. on Theory of Computing*, pp. 163–169, 1985.
12. F. F. Yao. Computational geometry. In J. van Leeuwen, ed., *Handbook of Theoretical Computer Science*, Volume A, Chapter 7. Elsevier, Amsterdam, 1990.
13. F. F. Yao, D. P. Dobkin, H. Edelsbrunner, and M. S. Paterson. Partitioning space for range queries. *SIAM J. Comput.*, 18(2):371–384, April 1989.

Received July 11, 1991, and in revised form April 9, 1992.