Some Unsolved Problems in Plane Geometry

A collection of simply stated problems that deserve equally simple solutions.

VICTOR KLEE
University of Washington
Seattle, WA 98195

If $S(t)$ is the number of mathematical problems that have been solved up to time $t$, and $U(t)$ is the number that have been explicitly considered but still remain unsolved, then probably $U(t)/S(t) \to \infty$ as $t \to \infty$. Yet most mathematical expositions concentrate on the denominator and ignore the much larger numerator. I like to redress the balance by talking or writing about unsolved problems, especially ones that have immediate intuitive appeal and can be understood with little background. The problems presented here are all concerned with the geometry of the Euclidean plane, which is almost as fertile a source of such problems as number theory and combinatorics. The problems are not new (one dates from 1916) but probably will be new to some readers. It will be interesting to see how many years pass before the problems are solved.

This paper is divided into two parts. In the first part, the problems are presented and some background information is provided. I hope this part will be accessible to all readers of the Magazine. The second part contains references and additional comments which are more advanced than the material of the first part. Nevertheless, it should be accessible to most readers.

I should emphasize that there is a tremendous number and variety of “elementary” unsolved problems in plane geometry. The small sample presented here consists of the ones that I find most appealing. This paper is an adaptation of [KH]. It is dedicated to Hugo Hadwiger for his seventieth birthday and to Paul Erdős for his sixty-fifth. The writings of Hadwiger and of Erdős are ideal sources of additional problems.

Problems

1. A colorful problem

With only two colors, you can paint an entire line so that no two points at unit distance receive the same color. Simply use red for each half-open interval $[n, n + 1]$ when $n$ is even, and white when $n$ is odd. That’s easy to do mathematically. However, with real paint it would be difficult to paint half-open rather than closed unit intervals, and to do the latter would produce two points at unit distance with the same color. Even with real paint, three colors suffice. For each integer $n$, use red for the half-open interval $[6n/3, (6n + 2)/3]$, white for $[(6n + 2)/3, (6n + 4)/3]$, and blue for $[(6n + 4)/3, (6n + 6)/3]$. A little sloppiness at ends of the interval doesn’t matter, because no number in the interval $[2/3, 4/3]$ is realized as the distance between two points that receive the same color.

What happens in the Euclidean plane $E^2$? That is:
(A) If all the points of the plane are to be painted so that no two points at unit distance receive the same color, what is the minimum number \( c \) of colors that can be used?

Though the exact value of \( c \) is unknown, it's easy to see \( 4 < c < 7 \). To prove \( c > 4 \), we show three colors are not sufficient, and to prove \( c < 7 \), we tell how to paint the plane in seven colors so that no two points at unit distance receive the same color.

Suppose the plane is painted in three colors—say red, white, and blue—so that no two points at unit distance receive the same color. For an arbitrary red point \( r \), consider the configuration shown in Figure 1. The triangles are equilateral with side length 1, so the other two vertices of \( r \)'s triangle are colored differently from \( r \) and from each other. Thus one is white and one blue, whence \( r' \) is red. By rotating the rhombus about \( r \), we obtain an entire circle of red points \( r' \). The circle includes two points at unit distance, and the resulting contradiction shows \( c > 4 \).

For an upper bound on \( c \), consider the tessellation of the plane by regular hexagons of side length \( 2/5 \). Paint one hexagon with color 7, and its six neighbors with colors 1 through 6. As the entire plane is tessellated by translates of this configuration, we can simply extend the coloring “by translation” (see Figure 2) and it turns out that no two points at unit distance receive the same color. Thus \( c < 7 \). We know, then, that \( c \) has one of the four values 4, 5, 6 and 7. But which one?

![Figure 1](image1.png)

![Figure 2](image2.png)

2. A modern way of squaring the circle?

The ancient Greek problem required the construction, for a given circle \( C \), of a square of the same area as \( C \). “Construction” meant that, starting from a line segment whose length is equal to the radius of \( C \), one should use ruler and compass to produce a side of a square whose area is that of \( C \). Starting from a segment of unit length, one of length \( \pi^{1/2} \) would be produced. As is now well known, that can’t be done with ruler and compass. However, the problem is unsolved (rather than provably unsolvable) for some other interpretations of “construction”.

Two subsets \( X \) and \( Y \) of a Euclidean space are said to be equivalent by finite decomposition if \( X \) can be partitioned into a finite number of sets \( X_1, \ldots, X_n \) and \( Y \) into sets \( Y_1, \ldots, Y_n \) such that \( X_i \) is congruent to \( Y_i \) for \( 1 \leq i \leq n \). Though now more than a half-century old, the following is sometimes called the “modern form of the problem of squaring the circle”:

(B) Can a circular region and a square region be equivalent by finite decomposition?
It is known that if a circular and a square region are equivalent then they have the same area, even though the individual sets in the partitions are not required to be “nice” or even to have area in the usual sense. Thus, the situation in $E^2$ contrasts sharply with that in higher dimensions. In $E^d$ for $d > 3$, any two sets that are bounded and have interior points are equivalent by finite decomposition. This is the Banach-Tarski “paradox”, which is discussed briefly in the second half of the paper. It provides a striking illustration of the fact that when we make a mathematical definition based on the apparent behavior of familiar objects, we may inadvertently create mathematical “monsters” whose behavior defies intuition.

3. Equichordal points

A chord of a plane region $R$ is a segment that joins two boundary points of $R$. A point $p$ of $R$ is an equichordal point if all chords through it are of the same length. For example, the center of a circle is an equichordal point. You might think that only a circular region could have an equichordal point, but that's far from the case. Consider a drawing arm as shown in Figure 3, with a marker at each end of the arm. The arm pivots at the point $p$ and is free to move back and forth through $p$. If we start with the ends of the arm at points $a$ and $b$ and move one end along a “reasonable” arc from $a$ to $b$ (one that doesn’t get too far from $p$ and that satisfies some other mild conditions), then the two ends together will draw the boundary of a region having $p$ as an equichordal point. In particular, as shown in Figure 3, a noncircular convex region may have an equichordal point. (Recall that a region is convex if it contains each segment joining two of its points.)

It is known that a plane region cannot have three equichordal points, but the following problem was posed in 1916 and is still unsettled:

(C) Can a plane convex region have two equichordal points?

The problem is also open for nonconvex regions, though in that case some fine points in the definition of “region” and “equichordal point” become relevant. Details are given in the second half of this paper.

A point $p$ of a plane region is an equichordal point if and only if the sum of distances $\delta(x,p) + \delta(y,p)$ is constant for all chords $[x,y]$ through $p$. Analogously, call $p$ an equireciprocal point if the sum $\delta(x,p)^{-1} + \delta(y,p)^{-1}$ is constant. The foci of an ellipse are equireciprocal points, but the following problem is open:

(D) Can a nonelliptical plane convex region have two equireciprocal points?

4. A tale of two problems

In considering the problems of this paper, it is natural to wonder whether anyone has a reasonable chance of solving them. I can’t answer that, except to say that problems of this sort are great equalizers among mathematicians, for solutions usually depend on clever ideas rather than extensive knowledge of theory or development of complicated mathematical machinery.

If age is a reliable guide, the equichordal problem should be the most difficult one considered here, for it is the oldest. However, unsolved problems are by their nature unpredictable, and that may be especially true of ones that are easily stated. By way of illustration, let’s consider two problems from the last century that long appeared to be very difficult despite their elementary nature.

The famous four-color problem, which originated in 1852, asks whether every planar map can be painted in four colors so that no two countries with a common boundary arc receive the same color. The problem was first mentioned in print in 1878, and the sufficiency of five colors was soon demonstrated. However, for decades, it was unknown whether any planar map really required five colors. This and Fermat’s problem were the two most famous unsolved problems in mathematics. Finally, in 1976, it was proved that four colors suffice. A considerable amount
of mathematical machinery was involved, but much of it had been available since 1880. The proof combined refinements of the machinery with many hours of electronic computation. Thus, the four-color problem has been solved, but there is still no easy solution.

The following problem was posed in 1893: If $S$ is a finite set of points in the plane, not all collinear, must there exist an ordinary line for $S$? (An ordinary line is one that contains exactly two points of $S$.) The problem still appeared difficult forty years later, and was sometimes mentioned along with the four-color problem as an apparently impossible though easily stated problem. Yet now, thanks to a clever idea, it appears to be almost trivial. The solution is given in the next paragraph.

Since the points of $S$ are not all collinear, it is possible to choose a point $p$ of $S$ and a line $L$ that misses $p$ but contains at least two other points of $S$. Since $S$ is finite, there are only finitely many such choices and hence there is one for which the distance from $p$ to $L$ is minimum. For any such minimizing choice, the line $L$ is ordinary. Otherwise, at least three points of $S$ are on $L$ and at least two of them are on the same side of the foot $f$ of the perpendicular from $p$ to $L$. If $x$ and $y$ are two such points and $x$ is closer to $f$, the distance from $x$ to the line $py$ is less than the distance from $p$ to the line $L=xy$. That contradicts the minimizing property of the pair $(p, L)$. (See Figure 4)

It will be interesting to see whether any of the open problems from this paper is eventually solved in such an elegant manner. In working on the problems, one may hope to find the clever idea that leads to such a solution. However, there is also the danger that after one has spent months in a futile attack on a problem, someone else may find a short and elegant solution!

5. Reflections on reflections

Suppose you live in a 2-dimensional room whose walls form a simple closed polygon, and each wall is a mirror. Is there necessarily a point from which a single light source will illuminate the entire room, using reflected rays as well as direct rays? Must every point of the room have this property? More briefly:

(E) Is every polygonal plane region illuminable from at least one of its points?
From each of its points?

These problems are unsolved, but there are nonpolygonal regions for which the answers are negative. One is shown in Figure 5. The upper elliptical arc has foci at $p$ and $p'$, the lower one at $q$ and $q'$. The broken lines are the major axes of the ellipses. You'll recall that a light ray issuing from one focus of an ellipse is immediately reflected through the other focus; from this it
follows that any ray crossing the major axis between the foci is immediately reflected between the foci. Thus, in the figure, a ray that issues from any point above (resp. below) the lower axis cannot reach the areas marked B (resp. A). Though this smooth, nonpolygonal region R is not illuminable from any point, the problem remains open for polygonal regions, even for those very close to the region of Figure 5.

A convex region is of course illuminable from each of its points, even without using reflections. But what happens when attention is restricted to a single ray rather than (as in the illumination problem) all rays issuing from a particular point? Of course, a single ray cannot cover all the points of the region, but it may come arbitrarily close to each point—in topological parlance, it may be dense. For example, if the region is the square shown in Figure 6, then every ray with rational slope is closed, like the one shown in the figure, but every ray with irrational slope is dense. There are many unsolved problems concerning the behavior of individual light rays in convex regions, though they are usually stated in terms of billiard ball paths rather than light rays. Perhaps the simplest is the following:

(F) Does every triangular region admit a dense light ray?

An affirmative answer is known for triangles (in fact, for closed polygons) in which each angle measure is a rational multiple of π.

6. Forming convex polygons

A set of points in the plane is in **general position** if no three of them are collinear. A much stronger requirement is that the points are the vertices of a convex polygon. For example, the set of 10 points in Figure 7 is in general position, but it does not contain the vertex-set of any convex k-gon for k > 6.

(G) For n ≥ 3, what is the smallest number f(n) such that whenever S is a set of more than f(n) points in general position in the plane, S contains the vertex-set of a convex n-gon?

It is known that f(3) = 2, f(4) = 4 and f(5) = 8. For larger n, it is not intuitively obvious that any size of S is sufficient to guarantee that S contains the vertex-set of a convex n-gon. However, it has been proved that

\[ 2^{n-2} \leq f(n) \leq \binom{2n-4}{n-2} \]  

(6.1)

and conjectured that \( f(n) = 2^{n-2} \).
7. Pushing disks around

The next problem deals with congruent circular disks in the plane and with the area that they cover. When two disks are far apart, the area of their union is simply the sum of the individual areas, but as they are pushed closer together, some overlapping occurs and the area of the union decreases. For two disks, the area of the union certainly cannot be increased by pushing the disks closer together. But what happens with three or more disks? Suppose their centers are tautly connected by inelastic string, as shown in Figure 8, and we are permitted to move the disks but not to break the string. Thus, we can decrease the distances between the centers but not increase them. Is it possible to increase the total area covered? A negative answer is expected and this is known for three disks. However, the problem is unsettled for any larger number of disks. Here is a formal statement:

(H) If \( C_1, \ldots, C_n \) and \( D_1, \ldots, D_n \) are congruent circular disks in the plane, centered at points \( p_1, \ldots, p_n \) and \( q_1, \ldots, q_n \) respectively, and if \( d(q_i, q_j) < \delta(p_i, p_j) \) for all \( i \) and \( j \), does it follow that

\[
\text{area of } \bigcup_i D_i < \text{area of } \bigcup_i C_i.
\]  

(7.1)

It has been proved that the area covered by the \( D_i \)'s is no more than 9 times the area covered by the \( C_i \)'s, but that's far from solving the problem. To understand the difficulty, note that if \( \alpha \) denotes area then

\[
\alpha(C_1 \cup C_2 \cup \cdots \cup C_n) = A_1 - A_2 + A_3 - \cdots + (-1)^{n-1} A_n,
\]

where \( A_1 \) is the sum of the individual areas and \( A_k \) is the sum of the areas of the \( k \)-at-a-time intersections. That is,

\[
A_k = \sum_{1 < j_1 < j_2 < \cdots < j_k < n} \alpha(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}).
\]

In particular, \( A_n \) is the area common to all \( n \) disks. As the disks are pushed closer together, \( A_1 \) is unchanged and \( A_2 \) does not decrease, so the two-term approximation \( A_1 - A_2 \) of \( \alpha(C_1 \cup \cdots \cup C_n) \) does not increase. However, the problem is complicated by the presence of \( A_3, \ldots, A_n \).

8. Pushing points around

There are many appealing problems that involve a mixture of plane geometry and number theory. My favorite is presented here.

Let us say that a set is rational (resp. integral) if only rational numbers (resp. integers) are realized as distances between points of the set. Each line in Euclidean \( d \)-space \( E^d \) contains a dense rational set, and it is known that each infinite integral set in \( E^d \) is actually contained in a
line. But what can be said about rational sets in the plane? In particular:

(I) Can every finite subset $X = \{x_1, \ldots, x_n\}$ of the plane be closely approximated by a rational set? If so, does the plane actually contain a dense rational set?

By closely approximated, we mean that for each $\varepsilon > 0$ there exist points $y_1, \ldots, y_n$ such that the set $Y = \{y_1, \ldots, y_n\}$ is rational and $\delta(x_i, y_i) < \varepsilon$ for all $i$. In other words, we would like to push the points $x_i$ to nearby positions $y_i$ in order to obtain a rational set.

To gain some insight into the problem, consider the case of three noncollinear points $x_1$, $x_2$ and $x_3$. Of course,

$$\delta(x_i, x_k) < \delta(x_i, x_j) + \delta(x_j, x_k) \tag{8.1}$$

for each permutation $(i, j, k)$ of $(1, 2, 3)$. Taking $y_1 = x_1$, we may move $x_2$ to a nearby position $y_2$ so as to preserve the inequalities (8.1) and make $\delta(y_1, y_2)$ rational. Then if $\rho_1$ and $\rho_2$ are rational numbers sufficiently close to $\delta(y_1, x_3)$ and $\delta(y_2, x_3)$ respectively, and if $C_i$ is the circle centered at $y_i$ with radius $\rho_i$, the intersection $C_1 \cap C_2$ will consist of two points, one of which is very close to $x_3$. (See Figure 9.) If $x_3$ is replaced by this nearby point $y_3$, a close rational approximation $\{y_1, y_2, y_3\}$ of the set $\{x_1, x_2, x_3\}$ is obtained. Thus for $n = 3$ the desired rational approximation can be obtained by simply moving the points one at a time. However, this simple approach doesn't work when $n > 4$ for there are $\binom{n}{2}$ point-pairs whose distances must be controlled.

9. Inscribed squares

The next unsolved problem is my favorite among those that involve a mixture of plane geometry and topology.

We say that a square $S$ is inscribed in a set $X$ if $X$ contains all four vertices of $S$. It doesn't matter how the edges of $S$ are related to $X$. For example, the square indicated in Figure 10 is inscribed in the set $C$ shown there. When can a plane set be expected to admit an inscribed square? Since any set with an interior point admits many such squares, we're really interested only in sets that are “thin” in some sense. The best that can be expected is an affirmative answer to the following question:

(J) Does every plane simple closed curve $C$ admit an inscribed square?

Recall that a simple closed curve is a set that is topologically the same as a circle. It need not be very “simple” at all from the viewpoint of Euclidean geometry. For example, the set $C$ of Figure 10 is a simple closed curve. One naturally tries certain continuity arguments in attempting to answer (J) affirmatively, and they are easily made rigorous when $C$ is the boundary of a convex region. Several authors have claimed incorrectly to extend them to an arbitrary simple closed curve $C$, but the extension has been made correctly only when $C$ is sufficiently smooth.

![Figure 9](image1)

![Figure 10](image2)
1. A colorful problem

For each positive integer $d$, let $A_d(d)$ denote the smallest $c$ such that all of Euclidean $d$-space $E^d$ can be painted with $c$ colors so that no two points at unit distance receive the same color. It follows from an observation of de Bruijn and Erdős [DE] that $c$ is unchanged if "all" is replaced by "every finite subset." A short proof can be based on Tychonov's theorem asserting that the product of compact topological spaces is compact. For let $P$ denote the set of all functions on $E^d$ to $[1, \ldots, c]$, topologized as the compact product $\times_{y \in E^d} Q_y$ where each factor $Q_y$ is equal to $[1, \ldots, c]$. For each finite $S \subseteq E^d$, let $P_S$ denote the closed subset of $P$ consisting of all $f \in P$ such that $f(a) \neq f(b)$ whenever $a$ and $b$ are points of $S$ at unit distance. If every finite subset of $E^d$ can be colored in $c$ colors so that no two points at unit distance receive the same color, the family of closed sets ($P_S$: finite $S \subseteq E^d$) has the finite intersection property and hence by compactness has nonempty intersection. Each point $f$ of the intersection yields a painting of all of $E^d$ in $c$ colors so that no two points at unit distance receive the same color.

The emphasis on unit distance has been mainly for conciseness. Note that $A_d(d)$ is the smallest $c$ such that whenever $E^d$ is covered by fewer than $c$ sets, then for each $\delta > 0$, at least one of the sets includes two points at distance $\delta$. A related number is $B_d(d)$, the smallest $c$ such that whenever $E^d$ is covered by fewer than $c$ sets, at least one of them is a $\Delta$-set, meaning that it realizes all positive numbers as distances between its point-pairs. The numbers $A_d(d)$ and $B_d(d)$ are similarly defined, except that for $i = 1$ they refer to coverings by mutually congruent sets, for $i = 2$ by closed sets, and for $i = 3$ by sets that are both closed and congruent. Plainly $B_d(d) \leq A_d(d),$ $A_0(d) \leq A_1(d) \leq A_3(d)$, and $B_0(d) \leq B_1(d) \leq B_3(d)$.

Problem (A), which was first discussed by Gardner [Ga] in 1960 and Hadwiger [Ha7] in 1961, asks for the determination of $A_0(2)$. It is not hard to see that $A_0(1)$ and $B_0(1)$ are equal to 2 for $i = 0, 1$ and to 3 for $i = 2, 3$. The construction based on the hexagonal tessellation of the plane shows that $A_3(2) < 7$, for it provides a covering of the plane by 7 mutually congruent closed sets such that the entire interval $[4/5, 2\sqrt{7}/5]$ is omitted from the distances determined by the point-pairs of any of the sets. In [Ha1, Ha2], Hadwiger showed for all $d$ that $B_2(d) > d + 2$ and $B_3(d) > 4d^2$, and Raiskii [Ra] and Woodall [Wo] improved the first result to $B_0(d) > d + 2$. For all $d > 5$, a deep study of "configurations" by Larman and Rogers [LR] and Larman [La'] led to further improvements. In particular, [LR] proved that $B_0(d) > d(d - 1)/6$ and [La'] that $B_0(d) > (n - 2)(n - 3)(n - 4)/178200$. Both of these papers contain additional results and conjectures of great interest. In particular, [La'] conjectures $B_0(d) > \frac{1}{4}(\frac{4}{3})^{d/4}$. In a recent letter, Paul Erdős conjectures the existence of $\delta > 0$ such that $A_0(d) > (1 + \delta)^d$ for all $d$, and he reports Peter Frankl's theorem that for some $k > 1$, $A_0(d) > d^k$ for all sufficiently large $d$.

2. A modern way of squaring the circle?

Problem (B), posed by Tarski [Ta'2] in 1925, was motivated by the theorem [Ta'1] that two polygonal plane regions are equivalent by finite decomposition if and only if they are equal in area. In this notion of equivalence, the pieces into which a set is decomposed are individually unrestricted, but they must be pairwise disjoint. For equivalence based on decompositions into polygonal regions that may intersect but not overlap, the theorem is due to J. Bolyai (1802-1860), a founder of noneuclidean geometry. For Bolyai's decomposition theory and higher-dimensional analogues, see Hadwiger [Ha5], Boltyanski [Bo1, 2] and Meschkowski [Me']. See Dubins, Hirsch and Karush [DHK] and Sallee [Sa'] for results and unsolved problems related to Tarski's problem.

When the pieces are unrestricted, it seems conceivable that two plane regions could be equivalent by finite decomposition even though they are unequal in area. However, this is
excluded by the result of Banach [Ba1] and Morse [Mo'] that the area function of plane geometry (and even 2-dimensional Lebesgue measure) can be extended to a function \( \mu \) that is:

(a) defined for all bounded subsets of \( E^2 \);
(b) finitely additive—\( \mu(A \cup B) = \mu(A) + \mu(B) \) whenever \( A \) and \( B \) are disjoint bounded subsets of \( E^2 \);
(c) invariant under rigid motions—\( \mu(A) = \mu(B) \) whenever \( A \) and \( B \) are congruent bounded subsets of \( E^2 \).

Though the Banach-Tarski paradox seems incredible, one might imagine that the number of sets in "paradoxical" decompositions is so large that our intuition is unreliable in dealing with such numbers. However, that is not the correct explanation. Robinson [Ro] showed that if \( X \), \( Y \) and \( Z \) are congruent spherical balls in \( E^3 \), then \( X \) can be partitioned into five sets \( X_1, \ldots, X_5 \) in such a way that the pieces \( X_1 \) and \( X_2 \) can be reassembled (moved by suitable rigid motions) to form \( Y \) and the pieces \( X_3, X_4 \) and \( X_5 \), one of which consists of a single point, can be reassembled to form \( Z \). It is hard to think of a more surprising result! The failure of our intuition does not stem from the number of pieces \( X_i \), but from the complicated nature of some of the pieces. In particular, the volume function of solid geometry cannot be extended to a function \( \mu \) that satisfies conditions (a), (b) and (c) with \( E^2 \) replaced by \( E^3 \). If the extension \( \mu \) is to be defined for all bounded subsets of \( E^3 \), finite additivity or invariance under rigid motions must be sacrificed. On the other hand, finite additivity and invariance under translations can be preserved. The crux of the matter is the fact that the group of translations in \( E^3 \) (or rigid motions in \( E^2 \)) is solvable, while the group of rigid motions in \( E^3 \) is not. See Sierpinski [Si], Dekker [De] and Meschkowski [Me'] for expositions of the Banach-Tarski paradox and related unsolved problems.

The Banach-Tarski paradox depends heavily on the axiom of choice, but there are other weird decompositions that are quite elementary and constructive in nature and work even in the plane. Some are described by Hadwiger, Debrunner, and Klee [HDK]. Others can be derived from Banach's easily proved theorem [Ba2] that for any two sets \( X \) and \( Y \) and one-to-one mappings \( f: X \to Y \) and \( g: Y \to X \) there are partitions \( X = X_1 \cup X_2 \) and \( Y = Y_1 \cup Y_2 \) for which \( fX_1 = Y_1 \) and \( gY_2 = X_2 \). For example, let us say that two subsets \( U \) and \( V \) of a normed linear space \( E \) are homothetic if each can be obtained from the other by a dilatation—equivalently, if \( U = p + \lambda V \) for some \( p \in E \) and \( \lambda > 0 \). An immediate consequence of Banach's theorem is that if \( X \) and \( Y \) are subsets of \( E \) that are bounded and have nonempty interior, then there are partitions \( X = X_1 \cup X_2 \) and \( Y = Y_1 \cup Y_2 \) such that \( X_i \) and \( Y_i \) are homothetic for \( i = 1, 2 \). In view of this fact, and of Tarski's result for polygonal regions [Ta'1], it should perhaps not be surprising if there is "a modern way of squaring the circle."

3. Equichordal points

The equichordal problem was raised by Fujiwara [Fu] in 1916 and by Blaschke, Rothe and Weitzenböck [BRW] in 1917. The general strategy of attack on the problem has been to assume the existence of a plane convex region \( R \) with two equichordal points \( p \) and \( q \), and then to derive many necessary properties of \( R \). A sufficiently long list might show how to construct \( R \), or might include two mutually contradictory properties and hence show \( R \) does not exist after all. For example, Wirtinger [Wi] showed that \( R \)'s boundary curve \( C \) is analytic, while an earlier author had claimed \( C \) could not be six times differentiable. That would have settled the problem but for errors in the earlier "proof." Note that if \( R \) does exist we could design a drawing instrument with two arms, pivoting at \( p \) and \( q \) respectively and hinged together at one end, so that as the hinge traces \( R \)'s boundary, the other ends of both arms do the same. On the basis of experiment, this seems to be impossible. However, a proof is lacking, and Petty and Crotty [PC] have described some noneuclidean plane geometries in which a convex body can have two equichordal points.

Let \( p \) and \( q \) be distinct points of \( E^2 \) and let \( L \) denote the line through \( p \) and \( q \). A construction of Hayashi [Ha'] and Hallstrom [Ha"] apparently produces a subset \( R \) of \( E^2 \) such that \( L \cap R \) is
a segment and each line through $p$ or $q$ intersects $R$ in a segment of the same length as $L \cap R$. However, it seems the set $R$ so constructed may be neither closed nor open nor convex. The ends of the mentioned segments (aside from $L \cap R$) lie on two curves, but those may oscillate wildly as they approach the line $L$. In particular, the boundary of $R$ may fail to be a simple closed curve and there may be segments in $L$ that are longer than $L \cap R$ and yet join two boundary points of $R$.

The deepest and most important paper on the equichordal problem is that of Wirting [Wi], whose results do not depend on convexity; instead, they require only that $R$‘s boundary should be a simple closed curve. See Hadwiger [Ha3] and Klee [K11] for expositions of the problem and lists of other references. Some recent inconclusive attacks on the equichordal problem have been made by Hallstrom [Ha"{a}] (who also considered the equireciprocal problem (D)) and McLachlan and Owens [MO].

Problem (D) was initially misstated in [K11], then correctly stated by Guy and Klee [GK].

4. A tale of two problems

Some accounts of the four-color conjecture attribute it to cartographers, some to A. F. Möbius. Neither attribution is supported by the careful historical research of May [Ma]. Apparently the conjecture was first formulated by Francis Guthrie, who studied mathematics at University College, London. His brother, Frederick, told the conjecture in 1852 to their teacher, A. DeMorgan, who communicated it to other mathematicians. The first published reference, in 1878, was associated with A. Cayley. The clever but erroneous solutions of Kempe [Ke] and Tait [Ta] appeared in 1879-80, and Kempe’s solution was accepted until the error was pointed out by Heawood [He] in 1890. Over the years, several other research mathematicians and scores of mathematical amateurs produced erroneous “proofs” of the four-color conjecture. Errors were so common that whenever a new “solution” appeared, mathematicians would automatically assume there must be a hole in it somewhere. Thus it was especially interesting that the final solution, by Appel and Haken in 1976, was based on the original idea of Kempe (see [AH1, 2, 3], [AHK], [Ha’]).

Among treatments of the four-color problem that appeared before it was solved, the monographs of Ore [Or] and Heesch [He’] and the article of Coxeter [Co3] are worthy of special mention. A popular account of the solution appears in [AH2], and post-solution monographs have been written by Saaty and Kainen [SK] and Barnette [Ba’].

The sphere is easily seen to be equivalent to the plane for the purposes of map coloring, and the problem may be considered on other surfaces as well. The **chromatic number** of a surface $S$ is the smallest number $c$ such that any map on $S$ can be colored with $c$ or fewer colors. Let $c_p$ denote the chromatic number of the surface obtained by adding $p$ handles to a sphere or, equivalently, drilling $p$ separate holes through a block of wood. The four-color theorem asserts $c_0 = 4$. Heawood [He] showed $c_p < H(p)$ for all $p > 0$, where $H(p)$ is the greatest integer not exceeding $(7 + \sqrt{1 + 48p})/2$. He also showed $c_1 = H(1)$, so that the chromatic number of a torus is 7. Over the years, other mathematicians established equality for other values of $p$, and finally in 1968 Ringel and Youngs [RY1, 2] were able to show $c_p = H(p)$ for all $p > 0$. It is interesting that this problem, dealing with all cases except $p = 0$, should have been settled before the four-color problem.

The problem on ordinary lines was posed by Sylvester [Sy] in 1893, rediscovered by Erdős in 1933, and solved by T. Gallai a few days later. The short proof given here is due to L. M. Kelly (reported by Coxeter [Col, 2]). A similarly short proof was given by Lang [La] for the dual result, asserting that if a finite family of lines in the plane is such that its members are not all parallel and do not pass through a common point, then there exists a point that lies on exactly two of the lines. See Motzkin [Mo] for a fuller discussion of the problem’s history.

The affirmative solution to Sylvester’s question says there is at least one ordinary line for any finite plane set whose $n$ points are not all collinear. But should there not be many such lines when $n$ is large? Dirac [Di] conjectured there are at least $[n/2]$ ordinary lines, and Kelly and
Moser [KM] proved there are at least \(3n/7\). For other results related to Sylvester’s problem and Dirac’s conjecture, and for additional references, see Crowe and McKee [CM], Chakerian [Ch], Grünbaum [Gr], Kelly and Rottenberg [KR], Meyer [Me"], and Erdős and Purdy [EP2]. For higher-dimensional extensions, see Motzkin [Mo], Bonnice and Kelly [BK], Edelstein [Ed] and Rottenberg [Ro]. This is only a small sample of the many papers that have been inspired by Sylvester’s problem.

5. Reflections on reflections

Problem (E) was stated by Klee [Kl2] but did not originate with him. Perhaps it was first posed by E. Straus in the early 1950’s. The nonilluminatable region related to the ellipse is \(P\). Ungar’s modification of an idea of Penrose and Penrose [PP]. By extending the idea it is possible to construct, for each positive integer \(k\), a plane region \(R_k\) with smooth boundary such that \(R_k\) is not illuminatable from any set of \(k\) points. This was noted in [KH], and Rauch [Ra’] obtained related results.

Problems (E) and (F) involve the mathematical “light rays” of geometric optics rather than the light waves that are closer to physical reality. (See Rauch and Taylor [RT] for behavior of the latter.) It is assumed that whenever a ray meets an inner point of an edge of the polygonal boundary, the angle of reflection is equal to the angle of incidence; and when it meets a vertex, it simply ends (or it may be assumed to return along its former path).

The stated result on rays (or billiard ball paths) in squares is due to König and Szücs [KS] (see also Hardy and Wright [HW]), and was first proved by using a theorem of Kronecker on simultaneous Diophantine approximation. A more elementary proof was given by Sudan [Su]. The result on the existence of dense light rays was proved in different ways by Zemlyakov and Katok [ZK] and Boldrighini, Keane and Marchetti [BKM]. For other references, problems and results on billiard ball problems, see these two papers, Portitsky [Po], Croft and Swinnerton-Dyer [CS], Schoenberg [Sc1,2], Halpern [Ha""'] and Kreinovic [Kr].

6. Forming convex polygons

The conjecture that \(f(n) = 2^{n-2}\) was made by Erdős and Szekeres [ES1] in 1935. They established the right side of (6.1) in [ES1], and the left side 25 years later in [ES2]. (See also Kalbfleisch, Kalbfleisch and Stanton [KKS].) For other problems similar in spirit to (G), see [ES1], a paper by Erdős and Purdy [EP1], and other papers referred to in these. Recently Erdős has asked: For \(n \geq 3\), what is the smallest number \(g(n)\) (if it exists) such that whenever \(S\) is a set of more than \(g(n)\) points in general position in the plane, \(S\) contains the vertex-set of a convex \(n\)-gon that has no points of \(S\) in its interior? Plainly \(g(3) = 2\) and \(g(4) = 4\). The existence of \(g(5)\) was proved independently by A. Ehrenfeucht and by H. Harborth [Ha"2], who showed \(g(5) = 9\). It is unknown whether \(g(6) \leq \infty\).

A tool of [ES1] is the fact that in any permutation of the integers \(1, \ldots, n^2 + 1\) there is a monotone subsequence of length \(n + 1\). A short proof of this was given by Seidenberg [Se]. (See also Chvátal and Komlós [CV] and Mirsky [Mi']).

7. Pushing disks around

Problem (H) is due to Thue Poulsen [TP] and Kneser [Kn], and was discussed also by Hadwiger [Ha4]. It was Kneser who established the relaxation of (H) which has an additional factor 9 on the right side, and a factor \(3^d\) for the analogous problem in \(E^d\). Hadwiger noted that the 1-dimensional analogue of (7.1) (dealing with congruent segments on a line) is valid, and attributed to Habicht and Kneser the fact that (7.1) holds when the centers \(p_i\) can be moved continuously to the centers \(q_i\) in such a way that, during the motion, the inter-center distances never increase. A proof of the latter result was published by Bollobás [Bo]. To understand the hypothesis of the Habicht-Kneser-Bollobás theorem, imagine that the \(p_i\)'s are connected by elastic bands; then the assumption is that the \(p_i\)'s can be moved to the positions \(q_i\) in such a way that no band stretches at any time.
Under the hypotheses of (H), does it follow that
\[ \text{area of } \cap \mathcal{D}_i > \text{area of } \cap \mathcal{C}_i? \]
This problem also appears to be open, though it is known at least that if the intersection of the \( \mathcal{C}_i \)'s is nonempty, then so is the intersection of the \( \mathcal{D}_i \)'s. The analogous result in \( E^d \) was established by Kirszbraun [Ki], and in recent years it has been extended and applied in remarkable ways. See Danzer, Grübaum, and Klee [DGK] for the history up to 1963, and see Minty [Mi1,2] for more recent developments.

8. Pushing points around
Anning and Erdős [AE] proved (i) every infinite integral set in \( E^2 \) is collinear, and (ii) \( E^2 \) contains a noncollinear integral set of \( n \) points for each \( n > 3 \). Erdős [Er] gave a simple proof of (i) for \( E^d \), and Steiger [St] extended (ii) to \( E^d \) by replacing 3 with \( d+1 \) and "noncollinear" with "not contained in any hyperplane."

Using an idea of Müller [Mü], Hadwiger and Debrunner constructed a dense rational set in the unit circle of \( E^2 \) (see [HDK]). This implies (ii) and raises the question of which plane sets can be closely approximated by rational sets. Perhaps

(a) there is a rational set that is dense in the plane,

a possibility mentioned by Hadwiger [Ha6] in attributing the problem to Erdős (who writes that he first heard it from Ulam in 1945 or 1946 (see [U1])). Perhaps (a) fails but

(b) every finite plane set can be closely approximated by a rational set,

which would settle a question attributed by Mordell [Mo] to I. J. Schoenberg. Or perhaps, as suggested by J. R. Isbell in [HDK], there is an integer \( n > 5 \) such that

(c) in any rational plane set of at least \( n \) points, some three points are collinear or some four are concyclic.

That would strongly negate (b) and hence (a). An example of Harborth [Ha"1] shows that (c) fails when \( n = 5 \), but the case of \( n = 6 \) appears to be unsettled. Problems analogous to those posed by (a) – (c) are open in \( E^d \) for all \( d > 2 \), and there are no obvious relationships among the problems in different dimensions.

As was remarked by Mordell, the problem of finding all rational sets of 4 points in the plane goes back to Brahmegupta, who was born in 598 A.D. The problem was in a sense solved by E. E. Kummer in 1848, though it is not obvious from his parametrization that every set of 4 points in \( E^2 \) is closely approximated by a rational set. The latter result was proved by Mordell [Mo] in 1959 and sharpened by Almering [Al] in 1963. Since then, several other results on rational sets have been discovered; most of them can be found through a paper by Ang, Daykin and Sheng [ADS] and its references. The difficulty of problem (I), and the fact that little progress has been made on it, is dramatically illustrated by the fact that known results provide no basis for choosing between the truth of (a) on the one hand and, on the other hand, the truth of (c) for \( n = 6 \).

9. Inscribed squares
For all sufficiently smooth simple closed curves in the plane, the existence of an inscribed square was proved by S'niřelman [Sn] and Jerrard [Je]. For the case in which the curve is the boundary of a convex region, elementary proofs were given by Emch [Em1,2], Zindler [Zi] and Christensen [Ch']. This case had been settled earlier by O. Toeplitz, but apparently his proof was never published (see Grünbaum [Gr]). Perhaps the general case of (J) could be settled by the methods of nonstandard analysis, in the spirit of Naren's proof [Na] of the Jordan curve theorem.

It is natural to wonder about higher-dimensional analogues of (J). One aspect was settled by Bielecki [Bi], who described a 3-dimensional convex body in which no rectangular parallelepiped can be inscribed. On the other hand, Pucci [Pu] showed that every 3-dimensional convex body admits an inscribed regular octahedron. The theorems of S'niřelman and Pucci were extended to \( E^d \) by Guggenheimer [Gu].
References

[Ba'] D. Barnette, A monograph on the four-color problem, to be published by the Mathematical Association of America.
[Bo'1] V. Boltianski, Equivalent and Equidecomposable Figures, Heath, Boston, 1963. (Translated by A. Henn and C. Watts from the first Russian edition, Moscow, 1956.)
[Ch'] C. Christensen, A square inscribed in a convex figure (Danish), Mat. Tidsskr, B 1950 (1950) 22–26.


[Ha'] W. Haken, An attempt to understand the four color problem, J. Graph Theory, 1 (1977) 193–206.


[He"'] H. Heesch, Untersuchungen zum Vierfarbenproblem, Bibliographisches Institut, Mannheim, 1969.


