LATTICE POINTS IN REGIONS

IVAN NIVEN\(^1\) AND HERBERT S. ZUCKERMAN\(^2\)

1. Let \(S\) be a bounded set of points in the Euclidean plane with a unit distance defined. If a rectangular coordinate system is imposed, a certain number of points of \(S\) are lattice points, i.e. points with integer coordinates. Let \(m(S)\) be the minimum number of lattice points of \(S\) under all possible choices of the axis system, and \(M(S)\) the maximum number. For example if \(S\) is a closed disk of diameter one, then \(m(S) = 0\) and \(M(S) = 2\). The definitions of \(m(S)\) and \(M(S)\) could be given in terms of a fixed rectangular coordinate system, with the set \(S\) being freely rotated and translated in the plane. It will be convenient in the proofs to use sometimes one and sometimes the other of these two formulations.

Although the definitions and theorems of this paper are given for 2-dimensional Euclidean space, the generalization to higher dimensions involves no difficulties whatsoever.

It is apparent that \(m(S) \leq M(S)\) for any set \(S\). R. M. Robinson suggested that the strict inequality holds for a nonempty bounded closed set, which is a more general result than we had formulated.

**Theorem 1.** If \(S\) is a nonempty bounded closed set then \(m(S) < M(S)\).

**Proof.** Select any points \(A\) and \(B\) with \(A \in S\) and \(B \notin S\). On the straight line segment \(AB\) let \(P\) be the point of \(S\) that is closest to \(B\); \(P\) exists because \(S\) is closed. Now impose a coordinate system with origin at \(P\). With this coordinate system suppose that \(S\) has \(r\) lattice points, so that \(m(S) \leq r \leq M(S)\). Let \(Q_1, Q_2, \ldots, Q_k\) be all the lattice points in the plane which are not members of \(S\) and each of which is within unit distance of some point of \(S\); this is a finite set of points because \(S\) is bounded. For each of these points \(Q_i\) there is a positive number \(\delta_i\) so that the disk of radius \(\delta_i\) with center at \(Q_i\) contains no point of \(S\). Define \(\delta\) as \(\min(\delta_1, \delta_2, \ldots, \delta_k)\).

With the set \(S\) held fixed, translate the coordinate system a distance \(\delta/2\) in the direction from \(P\) to \(B\). One lattice point at least is thereby removed from \(S\), and no lattice point is gained. With this new coordinate system, the set \(S\) has at most \(r - 1\) lattice points and so
2. **Measurable regions.** We turn now to a much more restricted class of point sets and establish the following result.

**Theorem 2.** If \( R \) is a bounded measurable region then \( m(R) \lesssim \mu(R) \lesssim M(R) \), where \( \mu(R) \) denotes the measure of \( R \).

If \( \mu(R) \) is not an integer the inequalities can of course be made strict.

**Proof.** Relative to any fixed coordinate system define the function

\[
\phi(x, y) = \begin{cases} 
1 & \text{if } (x, y) \in R, \\
0 & \text{if } (x, y) \notin R.
\end{cases}
\]

Define \( \psi(x, y) \) by

\[
\psi(x, y) = \begin{cases} 
1 & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1, \\
0 & \text{otherwise}.
\end{cases}
\]

If \( i \) and \( j \) are any integers then

\[
\int_0^1 \int_0^1 \phi(x + i, y + j) \, dx \, dy = \int_{-\infty}^\infty \int_{-\infty}^\infty \phi(x + i, y + j) \psi(x, y) \, dx \, dy
\]

\[
= \int_{-\infty}^\infty \int_{-\infty}^\infty \phi(x, y) \psi(x - i, y - j) \, dx \, dy
\]

\[
= \iint_{R \cap Q_{ij}} dx \, dy
\]

where \( Q_{ij} \) is the unit square \( i \leq x \leq i+1, j \leq y \leq j+1 \). Next define

\[
C(R, x, y) = \sum_{i,j} \phi(x + i, y + j)
\]

where the sum extends over all pairs of integers \( i, j \). The sum is finite since \( R \) is bounded. Thus \( C(R, x, y) \) counts the number of points \((x+i, y+j), i=0, \pm 1, \pm 2, \cdots, j=0, \pm 1, \pm 2, \cdots,\) that belong to \( R \). Thus \( C(R, 0, 0) \) is the number of lattice points belonging to \( R \). If we think of a new coordinate system parallel to the old one but with origin at \((x, y)\) we see that \( C(R, x, y) \) counts the number of lattice points belonging to \( R \) relative to the new coordinate system. Therefore we have

\[
m(R) \leq C(R, x, y) \leq M(R)
\]

for every pair of real numbers \( x, y \). Also by (1) we see that
\[ \int_0^1 \int_0^1 C(R, x, y) \, dx \, dy = \sum_{i,j} \int_0^1 \int_0^1 \phi(x + i, y + j) \, dx \, dy \]

(3)

\[ = \sum_{i,j} \int_R \int_R \, dy \, dx = \int_R \int_R \, dy \, dx = \mu(R). \]

This with (2) implies that \( m(R) \leq \mu(R) \leq M(R) \) because, for example,

\[ \int_0^1 \int_0^1 m(R) \, dx \, dy = m(R). \]

If we also assume that \( R \) is closed and nonempty, Theorem 2 can be strengthened.

**Theorem 3.** If \( R \) is a closed, nonempty, bounded, measurable region then \( m(R) \leq \mu(R) < M(R) \).

**Proof.** Impose a coordinate system so that \( R \) covers exactly \( m(R) \) lattice points, so that \( m(R) = C(R, 0, 0) \) in the notation of the proof of Theorem 2. Let \( \delta < 1 \) be a positive number so that for each lattice point \( P \) not in \( R \) the distance from \( P \) to every point of \( R \) exceeds \( \delta \).

It follows that \( C(R, x, y) = m(R) \) for all \((x, y)\) satisfying \( x^2 + y^2 \leq \delta^2 \). Let \( T_1 \) be the region defined by

(4) \[ 0 \leq x, \quad 0 \leq y, \quad x^2 + y^2 \leq \delta^2, \]

and \( T_2 \) the region

(5) \[ 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad x^2 + y^2 > \delta^2. \]

First we observe that

\[ \int \int_{T_1} C(R, x, y) \, dx \, dy = \int \int_{T_1} m(R) \, dx \, dy. \]

Also, since the two regions \( T_1 \) and \( T_2 \) comprise the unit square, \( 0 \leq x \leq 1, \quad 0 \leq y \leq 1 \), equation (3) gives

\[ \mu(R) = \int \int_{T_1} C(R, x, y) \, dx \, dy + \int \int_{T_2} C(R, x, y) \, dx \, dy \]

\[ \leq \int \int_{T_1} m(R) \, dx \, dy + \int \int_{T_2} M(R) \, dx \, dy \]

\[ = \pi \delta^2 m(R)/4 + M(R)(1 - \pi \delta^2/4) \]

\[ = M(R) - \{ M(R) - m(R) \} \pi \delta^2/4 < M(R) \]

by Theorem 1.
3. **Regions constructed of disks.** The inequalities of Theorems 2 and 3 can be made strict if we restrict the class of regions as follows. We begin with a measurable topological disk \( D \), i.e., the topological equivalent of a circle plus interior. Let \( D_1 \) be a measurable topological disk such that \( D_1 \subseteq D \) and such that the intersection of the boundaries of \( D_1 \) and \( D \) is at most a finite set of points.

Let \( T_1 \) be the closure of \( D - D_1 \). Let \( D_2 \) be a measurable topological disk such that \( D_2 \subseteq T_1 \) and such that the intersection of the boundaries of \( D_2 \) and \( T_1 \) is at most a finite set of points. Let \( T_2 \) be the closure of \( T_1 - D_2 \). Continuing, let \( D_3 \) be a measurable topological disk such that \( D_3 \subseteq T_2 \) and such that the intersection of the boundaries of \( D_3 \) and \( T_2 \) is at most a finite set of points. Let \( T_3 \) be the closure of \( T_2 - D_3 \). Repeat this procedure a finite number of times, obtaining the closed set \( T_j \). Say that an admissible region is one obtained by taking the union of a finite number of disjoint nonempty sets of the type \( T_j \).

We note that an admissible region satisfies the hypotheses of Theorem 3. Also, an admissible region has a well-defined boundary, namely the union of all the simple closed curves bounding the disks used in the construction. Except for a finite number of points of the boundary of an admissible region, the boundary in a sufficiently small neighborhood of a boundary point is an arc. The finite number of exceptional points will be called *junction* points.

**Theorem 4.** If \( R \) is an admissible region in the sense defined above, then \( m(R) < \mu(R) < M(R) \).

In view of Theorem 3 it suffices to prove \( m(R) < \mu(R) \). We begin the proof with two lemmas.

**Lemma 5.** If \( S \) is any finite set of points in the plane, and \( P_1 \) is a point not in \( S \), then a coordinate system can be imposed so that \( P_1 \) is a lattice point but no point of \( S \) is a lattice point.

**Proof.** Let a coordinate system be imposed so that \( P_1 \) is a lattice point. The set of points \( S \) can be separated into two disjoint sets \( S_1 \) and \( S_2 \), where the members of \( S_1 \) are lattice points but the members of \( S_2 \) are not. Let \( \delta \) be a positive real number smaller than the shortest distance of any point of \( S_2 \) to a lattice point. Rotate the coordinate system, with \( P_1 \) as center of rotation, through a sufficiently small angle so that no lattice point in the vicinity of any point of \( S_2 \) is moved by more than a distance \( \delta \), and so that the points of \( S_1 \) are no longer lattice points. Such a rotation gives a coordinate system satisfying the conditions of the lemma.
Lemma 6. If $R$ is an admissible region, a coordinate system can be imposed so that one and only one lattice point lies on the boundary of $R$, and it is not a junction point of the boundary of $R$.

Proof. Let $S$ be the set of junction points of $R$, and let $P_1$ be any point on the boundary of $R$, with $P_1$ not in $S$. Then by Lemma 5 a coordinate system $C_1$ can be imposed so that $P_1$ is a lattice point, and so that no point of $S$ is a lattice point. Suppose that all the lattice points on the boundary of $R$ are $P_1, P_2, \ldots, P_r$ with $r \geq 2$. None of these points is a junction point. Choose $\delta > 0$ so that except for the points $P_1, P_2, \ldots, P_r$, the distance from every lattice point to the boundary of $R$ exceeds $\delta$. We also choose $\delta$ to be less than the distance from any of $P_1, P_2, \ldots, P_r$ to any junction point on the boundary of $R$. We separate the proof into two cases, $r = 2$ and $r > 2$.

If $r = 2$ we try to remove $P_2$ from the boundary by a small rotation of the coordinate system with center of rotation $P_1$. This might not succeed because the boundary of $R$ in the vicinity of $P_2$ might be a circular arc with center $P_1$. If so, then starting from $C_1$ again we try to remove $P_1$ from the boundary of $R$ by a small rotation of the coordinate system with center of rotation $P_2$. If this does not remove $P_1$ from the boundary then we know that the boundary of $R$ in the neighborhoods of $P_1$ and $P_2$ consists of two circular arcs with centers at $P_2$ and $P_1$ respectively.

In this case we move the entire coordinate system from its original position $C_1$ along a path parallel to the circular arc at $P_1$, starting at $P_1$ and moving along the circular arc in one direction. In this motion of the coordinate system, no rotation is involved. This must remove the lattice point at $P_2$ from the boundary of $R$, since otherwise this would mean that the boundary of $R$ in the neighborhood of $P_2$ would consist of a small circular arc intersected by another. But this contradicts the fact that the local boundary of $R$ is a simple closed curve in the neighborhood of $P_2$. These motions of the coordinate system can be sufficiently small so that the total distance moved by any lattice point in the vicinity of $R$ is less than $\delta$, and so we get exactly one lattice point, $P_1$ or $P_2$, on the boundary. And this lattice point is not a junction point.

Next we turn to the other case of the proof, $r > 2$. Rotate the coordinate system, with $P_1$ as center of rotation, so that no lattice point in the vicinity of $R$ moves a greater distance than $\delta/2$. If in the process of rotation the lattice points $P_2, P_3, \ldots, P_r$ move off the boundary, the proof of the lemma is complete in this case. If not, we may presume that for a rotation about $P_1$ through some small
angle the points $P_2, \ldots, P_k$, with $k \leq r$, remain on the boundary of $R$. (As to notation here, the labels $P_2$ etc. stay with the lattice points as they move.) From the original position of the coordinate system, say $C_1$, we may presume that we have moved to a position $C_2$ with lattice points $P_1, P_2, \ldots, P_k$ on the boundary, and such that a small rotation of the coordinate system in both directions with $P_1$ as center of rotation does not move $P_2, \ldots, P_k$ off the boundary of $R$.

Next, starting from the coordinate system position $C_2$, make a small rotation with $P_2$ as center. This rotation will move $P_3, \ldots, P_k$ off the boundary for the following reason. If $P_3$ for example stayed on the boundary of the region $R$, this would mean that the local boundary in the neighborhood of $P_3$ would consist of a small circular arc with center at $P_1$ (relative to position of axes $C_1$) intersected by a small circular arc with center $P_2$ (relative to position of axes $C_2$). But the local boundary of $R$ at $P_3$ is a simple closed curve, so this is impossible. Thus there remain only two lattice points, $P_1$ and $P_2$, on the boundary of $R$. But this case has already been treated, so the proof of Lemma 6 is complete.

We now use Lemma 6 to prove Theorem 4. We impose a coordinate system so that exactly one lattice point, say $P$, lies on the boundary of $R$, and $P$ is not a junction point. Now the boundary of $R$ in the neighborhood of $P$ is an arc (of a disk) and so there are points $P_1$ interior to $R$ and $P_2$ exterior to $R$ such that the points of the straight line segment $PP_1$ belong to $R$, but the points of the straight line segment $PP_2$ (except $P$ itself) are not in $R$. By translating the coordinate system twice, once in the direction $PP_1$ and once in the direction $PP_2$, we can move this lattice point $P$ off the boundary, relocating it either inside $R$ or outside $R$. This can be done without disturbing the in-or-out relationship of all the lattice points in the plane other than $P$.

Denote the original coordinate system, with $P$ a lattice point on the boundary of $R$, by $\Gamma$. Denote the coordinate system resulting from a small translation of $\Gamma$ in the direction $PP_1$ by $\Gamma_1$, and similarly the system resulting from a small translation of $\Gamma$ in the direction $PP_2$ by $\Gamma_2$. Let $(x_2, y_2)$ be the coordinates of the origin of the $\Gamma_2$ system when related to the $\Gamma_1$ system. Let the function $C(R, x, y)$ as defined in the proof of Theorem 2 refer to the $\Gamma_1$ coordinate system. Then we have

$$C(R, 0, 0) = C(R, x_2, y_2) + 1 \geq m(R) + 1.$$  

Furthermore with the $\Gamma_1$ coordinate system there are no lattice points on the boundary of $R$, and there are $C(R, 0, 0)$ lattice points in the interior of $R$. Let $\delta > 0$ be chosen so that for every one of these
C(R, 0, 0) lattice points, a disk of radius \( \delta \) with center at the lattice point lies entirely inside \( R \). Then for each \((x, y)\) satisfying \( x^2 + y^2 \leq \delta^2 \) we see that \( C(R, x, y) = C(R, 0, 0) \). With \( T_1 \) and \( T_2 \) defined as in (4) and (5) we use (3) to get

\[
\mu(R) = \int \int_{T_1} C(R, x, y) \, dx \, dy + \int \int_{T_2} C(R, x, y) \, dx \, dy
\]

\[
= \int \int_{T_1} C(R, 0, 0) \, dx \, dy + \int \int_{T_2} C(R, x, y) \, dx \, dy
\]

\[
\geq \int \int_{T_1} \{m(R) + 1\} \, dx \, dy + \int \int_{T_2} m(R) \, dx \, dy
\]

\[
= m(R) + \int \int_{T_1} \, dx \, dy > m(R),
\]

as asserted in Theorem 4.

University of Oregon and
University of Washington