# Economical coverings of sets of lattice points

Noga Alon *

Department of Mathematics

Raymond and Beverly Sackler Faculty of Exact Sciences

Tel Aviv University, Tel Aviv, Israel

**Abstract**

Let $t(n, d)$ be the minimum number $t$ such that there are $t$ of the $n^d$ lattice points

$$\{(x_1, \ldots, x_d) : 1 \leq x_i \leq n\}$$

so that the $\binom{t}{2}$ lines that they determine cover all the above $n^d$ lattice points. We prove that for every integer $d \geq 2$ there are two positive constants $c_1 = c_1(d)$ and $c_2 = c_2(d)$ such that for every $n$

$$c_1 n^{d(d-1)/(2d-1)} \leq t(n, d) \leq c_2 n^{d(d-1)/(2d-1)} \log n.$$

The special case $d = 2$ settles a problem of Erdös and Purdy.

*Research supported in part by a United States Israel BSF Grant and by a Bergmann Memorial Grant

## 1 Introduction

We say that a set of points $S$ in an Euclidean space (of any dimension) *determines* a line $l$ if $l$ contains at least 2 points of $S$. Let $L = L(n, 2) = \{(x_1, x_2) : 1 \leq x_1, x_2 \leq n\}$ denote the set of all points in the $n$ by $n$ square lattice. A subset $S$ of $L$ is called an $(n, 2)$-*covering set* if the union of all the lines determined by $S$ contains all the points in $L$. Let $t(n, 2)$ denote the minimum cardinality of an $(n, 2)$-covering set.

Erdös and Purdy (see [1]) raised the problem of estimating $t(n, 2)$. They mentioned that it is not hard to see that $t(n, 2) \geq \Omega(n^{2/3})$ and asked if $t(n, 2) = o(n)$. In this short paper we answer this question affirmatively and show that $t(n, 2) \leq O(n^{2/3} \log n)$. Therefore, $t(n, 2)$ is indeed $o(n)$ and the gap between the upper and the lower bounds for this quantity is a $\log n$ factor.

Our proof combines simple geometric and probabilistic arguments with some of the standard techniques used in the study of Diophantine Approximations. The method works in dimensions bigger than 2 as well. Let $L(n, d)$ denote the set of all $n^d$ vectors $(x_1, \ldots, x_d)$, where $1 \leq x_i \leq n$ are integers for all $i$. A subset $S$ of $L(n, d)$ is an $(n, d)$-*covering set* if the lines determined by $S$ cover all points of $L(n, d)$. Let $t(n, d)$ denote the minimum possible cardinality of an $(n, d)$-covering set. Our main result is the following.

**Theorem 1.1** *For every integer $d \geq 2$ there are two positive constants $c_1 = c_1(d)$ and $c_2 = c_2(d)$ such that for every $n$:*

$$c_1 n^{d(d-1)/(2d-1)} \leq t(n, d) \leq c_2 n^{d(d-1)/(2d-1)} \log n.$$

## 2 The lower bound

In this section we prove the easy part of Theorem 1.1, by showing that for every $d$ there exists a positive constant $c_1 = c_1(d)$ such that $t(n, d)$ is at least $c_1 n^{d(d-1)/(2d-1)}$ for all $n$. Put $L = L(n, d)$. For a line $l$ in $R^d$, let $|l|$ denote the number of points of $L$ contained in $l$. Observe that if $|l| \geq 2$ then $l$ can be presented by its parametric equations $X_i = a_i + p_i z$, $(1 \leq i \leq d)$, where $a_i, p_i$ are integers and the greatest common divisor of $p_1, \ldots, p_d$ is 1. We say that a line $l$ with the above presentation is of *type* $q$ if $q = max|p_i|$. The *direction* of $l$ is the vector $(p_1, \ldots, p_d)$. Note that this vector and its inverse $(-p_1, \ldots, -p_d)$ represent the same direction. We need the following simple fact.

**Fact:** For each $q \geq 1$, the number of distinct directions of lines of type $q$ is at most $d(2q+1)^{d-1} \leq d(3q)^{d-1}$.

**Proof** There are $d$ possibilities for choosing an index $i$ such that $|p_i| = q$, and since a vector and its inverse define the same direction we may assume that $p_i = q$. Any other $p_j$ is an integer between $-q$ and $q$ implying the desired estimate. $\Box$

Let $S$ be a subset of cardinality $t$ of $L$. Let $F$ denote the set of all lines determined by $S$. Clearly, $|F| \leq \binom{t}{2}$. Let $P(S)$ denote the set of all the points in the lattice $L$ which are covered by the union of all the lines in $F$. Obviously,

$$|P(S)| \leq \sum_{l \in F} |l|. \tag{1}$$

Clearly, for each line $l$ of type $q$, $|l| \leq n/q$. Also, the number of lines in $F$ in each direction is at most $t/2$, since a set of $t$ points cannot determine more than $t/2$ parallel lines. Let $f_q$ denote the number of lines of type $q$ in $F$. By the Fact and the remark above $f_q \leq \frac{t}{2} d(3q)^{d-1}$. Therefore, the right hand side of (1) is at most the maximum possible value of $\sum_{q \geq 1} f_q \frac{n}{q}$, subject to the constraints $\sum_{q \geq 1} f_q \leq \binom{t}{2}$ and $f_q \leq \frac{t}{2} d(3q)^{d-1}$. This last maximum is obviously attained when $\sum_{q \geq 1} f_q = \binom{t}{2}$ and $f_q$ is as large as it can be for all $q < s$ and is 0 for all $q > s$, where $s$ is an appropriately chosen integer. Therefore, at the maximum, $f_q = \frac{t}{2} d(3q)^{d-1}$ for all $q < s$ and $f_q = 0$ for all $q > s$. Since $\sum_{q \geq 1} f_q = \binom{t}{2}$, a simple calculation shows that $s \leq 1 + t^{1/d}/3^{(d-1)/d}$. This implies that the above maximum is at most $\sum_{q=1}^{s} \frac{t}{2} d(3q)^{d-1} \frac{n}{q} < c_3(d) n t^{(2d-1)/d}$. We have thus proved that if $|S| = t$ then $|P(S)| \leq c_3(d) n t^{(2d-1)/d}$. Since if $S$ is an $(n, d)$-covering set then $P(S) = L$ (and hence $|P(S)| = n^d$), this gives the following:

**Lemma 2.1** *If $S$ is a subset of cardinality $t$ of $L(n, d)$ then the number of points of $L(n, d)$ covered by the lines determined by $S$ is at most $c_3 n t^{(2d-1)/d}$, where $c_3$ is a positive constant dependeing only on $d$. Thus, for every $d$ there is a positive constant $c_1 = c_1(d)$ such that $t(n, d) \geq c_1 n^{d(d-1)/(2d-1)}$ for all $n$.* $\Box$

## 3 The upper bound

In this section we prove the upper bound in Theorem 1.1. For convenience we omit the floor and ceiling signs and assume that the fractional powers of $n$ appearing here are all integers. Since we deal with fixed values of $d$ and large values of $n$, it is easy to see that this can indeed be assumed without loss of generality. We start with the following somewhat technical lemma.

**Lemma 3.1** *For every two lattice points $(x_1, \ldots x_d)$ and $(a_1, \ldots, a_d)$ in $L(n,d)$ there are $d$ integers $p_1, \ldots, p_d$ and a real number $z$, such that for $q = max\{|p_i| : 1 \le i \le d\}$:*

$$1 \le q \le n^{(d-1)/(2d-1)}, \tag{2}$$

$$|(x_i - a_i) - p_i z| \le \frac{n^{(2d-2)/(2d-1)}}{q} \quad for \quad all \quad 1 \le i \le d, \tag{3}$$

*and there exists an index $j$, $1 \le j \le d$, such that*

$$p_j = q \quad and \quad (x_j - a_j) - p_j z = 0. \tag{4}$$

**Proof** The proof is based on the standard argument of Dirichlet used in the study of approximation of reals by rationals, (see, e.g., [2]). Let us change the indices, if needed, so that

$$|x_d - a_d| = max\{|x_i - a_i| : 1 \le i \le d\}.$$

If this maximum is zero the result is trivial (since in this case we can take $p_i = 1$ for all $i$ and $z = 0$). Otherwise put $Q = n^{1/(2d-1)}$ and define $\alpha_i = \frac{x_i - a_i}{x_d - a_d}$ for $1 \le i \le d-1$. Consider the $(d-1)$-dimensional unit cube $0 \le y_i < 1$, $(1 \le i \le d-1)$ and split it into $Q^{d-1}$ identical subcubes by drawing hyperplanes parallel to its facets, where the distance between each pair of consecutive parallel hyperplanes is $1/Q = n^{-1/(2d-1)}$. For each integer $j$, $0 \le j \le Q^{d-1}$, let $P_j$ be the point $(j\alpha_1 (mod\, 1), \ldots, j\alpha_{d-1} (mod\, 1))$ in the above unit cube. Since there are $Q^{d-1} + 1$ such points, there are two of them, say $P_l$ and $P_m$, where $l < m$, that lie in the same subcube. Define $q = m - l$ and define $p_i$ by the equality $q\alpha_i = p_i + \epsilon_i$, where $p_i$ is an integer and $|\epsilon_i| \le 1/Q$.

Define $p_d = q$ and $z = (x_d - a_d)/q$. Since the absolute value of each $\alpha_i$ is at most 1 it follows that $|p_i| \le q \le Q^{d-1} = n^{(d-1)/(2d-1)}$ for all $1 \le i \le d$. In addition, for each $1 \le i \le d$

$$|q \frac{x_i - a_i}{x_d - a_d} - p_i| \le \frac{1}{Q},$$

implying

$$|(x_i - a_i) - p_i z| = |(x_i - a_i) - p_i \frac{x_d - a_d}{q}| \leq \frac{|x_d - a_d|}{qQ} \leq \frac{n^{(2d-2)/(2d-1)}}{q}.$$

Moreover, for $i = d$ we have $(x_d - a_d) - p_d z = 0$. This completes the proof of the lemma. □

Using the last lemma we prove the following:

**Lemma 3.2** *For every fixed $d \geq 2$ there are two positive constants $c_4 = c_4(d)$ and $c_5 = c_5(d)$ such that the following holds. For every $n \geq c_5$ and for every lattice point $(x_1, \ldots, x_d) \in L(n, d)$, there are at least $\frac{1}{2}n^d$ lattice points $(a_1, \ldots, a_d)$ in $L(n, d)$ satisfying*

$$(c_4 + 2)n^{(d-1)/(2d-1)} \leq a_i \leq n - (c_4 + 2)n^{(d-1)/(2d-1)} \quad \textit{for all} \quad 1 \leq i \leq d \tag{5}$$

*such that there exist $d$ integers $p_1, \ldots, p_d$ and a real number $z$ so that*

$$1 \leq |p_i| \leq n^{(d-1)/(2d-1)}, \tag{6}$$

*and*

$$|(x_i - a_i) - p_i z| \leq c_4 n^{(d-1)/(2d-1)} \quad \textit{for all} \quad 1 \leq i \leq d. \tag{7}$$

**Proof** We prove the lemma with $c_4 = 4d9^{d-1}$, where $c_5$ is chosen so that for $n \geq c_5$ the number of lattice points $(a_1, \ldots, a_d) \in L(n, d)$ violating (5) is at most $\frac{1}{4}n^d$.

By Lemma 3.1, for every $(a_1, \ldots, a_d)$ in $L(n, d)$ there are integers $p_1, \ldots p_d$ and a real number $z$ such that if $q$ is the maximum absolute value of the integers $p_i$, then inequalities (2) and (3) hold. Therefore, inequality (6) holds, by (2). If $q \geq \frac{1}{c_4}n^{(d-1)/(2d-1)}$ then inequality (7) follows from (3). Therefore, by the choice of $c_5$, in order to complete the proof of the lemma it suffices to show that the number of points $(a_1, \ldots, a_d)$ for which the above $q$ is smaller than $\frac{1}{c_4}n^{(d-1)/(2d-1)}$ is at most $\frac{1}{4}n^d$.

The number of these points can be estimated as follows: For each fixed value of $q$, $1 \leq q \leq \frac{1}{c_4}n^{(d-1)/(2d-1)}$ , there are $d$ possibilities for choosing the index $j$ for which (4) holds, and there are at most $n$ possibilities for choosing $a_j$. Once these choices are made, $z$ is determined. There are at most $(2q + 1)^{d-1} \leq (3q)^{d-1}$ choices for the other integers $p_j$. Given these choices, each of the numbers $x_i - a_i$ must fall into an interval of length $2n^{(2d-2)/(2d-1)}/q$. Obviously, this implies that there are at most $2n^{(2d-2)/(2d-1)}/q + 1 \leq 3n^{(2d-2)/(2d-1)}/q$ ways to choose each $a_i$, for all $i \neq j$.

The total number of choices is thus at most

$$\sum_{q=1}^{n^{(d-1)/(2d-1)}/c_4} dn(3q)^{d-1}(3\frac{n^{(2d-2)/(2d-1)}}{q})^{d-1}$$

$$\sum_{q=1}^{n^{(d-1)/(2d-1)}/c_4} 9^{d-1}dn^{1+\frac{(2d-2)(d-1)}{2d-1}}$$

$$= \frac{9^{d-1}d}{c_4}n^d = \frac{1}{4}n^d.$$

This completes the proof. □

For a lattice point $\mathbf{a} = (a_1, \ldots, a_d)$ that satisfies (5), let $B(\mathbf{a})$ denote the set of all points $(y_1, \ldots, y_d) \in L(n, d)$ that satisfy

$$|a_i - y_i| \leq (c_4 + 2)n^{(d-1)/(2d-1)} \quad \textit{for all} \quad 1 \leq i \leq d,$$

where $c_4 = c_4(d)$ is the constant appearing in Lemma 3.2. The last ingredient we need for the proof of the upper bound in Theorem 1.1 is the following simple lemma.

**Lemma 3.3** *Let $(x_1, \ldots, x_d)$ be a latice point in $L(n, d)$, and suppose that the lattice point $\mathbf{a} = (a_1, \ldots, a_d)$ satisfies (5). Suppose, further, that there are $d$ integers $p_1, \ldots, p_d$ and a real number $z$ such that (6) and (7) hold. Then, there is a line containing $(x_1, \ldots x_d)$ that contains at least two points of $B(\mathbf{a})$.*

**Proof** Let $B$ denote the convex hull of the points in $B(\mathbf{a})$. Observe that $B$ is a $d$-dimesnional cube with side length $(2c_4 + 4)n^{(d-1)/(2d-1)}$. Consider the line $l$ given by

$$X_i = x_i - p_i w, \quad (1 \leq i \leq d, -\infty < w < \infty).$$

This line contains the point $(x_1, \ldots, x_d)$ (corresponding to $w = 0$). Let $(v_1, \ldots, v_d)$ be the point corresponding to $w = z$ in this line, (i.e., $v_i = x_i - p_i z$ for $1 \leq i \leq d$). By (7) this point lies inside the cube $B$. Moreover, its distance from each of the facets of $B$ is at least $2n^{(d-1)/(2d-1)}$. Since every integral value of the parameter $w$ gives a lattice point on $l$, this, together with the fact that the integers $p_i$ satisfy (6) imply that the lattice points on $l$ corresponding to, say, $w = \lceil z \rceil$ and to $w = \lceil z \rceil + 1$ both belong to $B(\mathbf{a})$, completing the proof. □

**Proof of Theorem 1.1** The lower bound is proved in Lemma 2.1. To prove the upper bound, fix $d \geq 2$ and let $c_4 = c_4(d)$ and $c_5 = c_5(d)$ be as in Lemma 3.2. Suppose $n \geq c_5$ and choose randomly and independently (with repetitions), $\lceil d \log n \rceil + 1$ lattice points $(a_1, \ldots, a_d)$ satisfying (5), where each point is chosen independently according to a uniform distribution on the points satisfying (5). Let $A$ be the random set of all chosen points and let $S = S(A)$ be the union of all the sets $B(\mathbf{a})$, where $a \in A$. Clearly

$$|S| < (\lceil d \log n \rceil + 1)(2c_4 + 5)^d n^{d(d-1)/(2d-1)} \leq c_6 n^{d(d-1)/(2d-1)} \log n,$$

where $c_6 = c_6(d)$ is a constant depending only on $d$.

By Lemma 3.2, for every fixed lattice point $(x_1, \ldots, x_d)$ in $L(n, d)$, the probability that there is **no** point $\mathbf{a} = (a_1, \ldots, a_d)$ among the chosen ones so that the assertion of the lemma holds for this point is at most $(\frac{1}{2})^{\lceil d \log n \rceil + 1} < 1/n^d$. Hence, the expected number of points $(x_1, \ldots, x_d)$ for which there is no such $\mathbf{a} \in A$ is less than 1 and hence there is a choice for $A$ such that there are no such points.

However, by Lemma 3.3, for such an $A$, $S(A)$ is an $(n, d)$ covering set. We have thus proved that if $n \geq c_5(d)$ then $t(n, d) \leq c_6(d) n^{d(d-1)/(2d-1)} \log n$, and this clearly implies the upper bound in Theorem 1.1 with, e.g., $c_2(d) = max\{c_6(d), c_5(d)\}$. $\square$

It would be interesting to decide if the $\log n$ factor in the upper bound in Theorem 1.1 is necessary.

## References


[1] Richard K. Guy, *Unsolved Problems in Number Theory*, Springer Verlag, New York and Heidelberg, 1981, p. 133.

[2] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, (Forth Edition), Oxford University Press, London, 1959, Chapter XI.