On Distance Sets in the Triangular Lattice

Tanbir Ahmed
Department of Computer Science and Software Engineering
Concordia University, Montréal, Canada
ta_ahmed@cs.concordia.ca

David Jacob Wildstrom
Department of Mathematics
University of Louisville, Louisville, KY 40292
dwildstr@erdos.math.louisville.edu

Abstract

In this note, we investigate a problem on maximum planar distance sets by Erdős and Fishburn [2], and prove a recent conjecture presented by Ahmed and Snevily [1] on distance sets in the triangular lattice. We also enumerate the number of distinct distances in a hexagonal array in the traingular lattice.

1 Introduction

A point set X in the Euclidean plane is called a k-distance set if there are exactly k different distances between distinct points in X. Let $g(k)$ denote the maximum number of points in a k-distance set. For example, it is well known that $g(1) = 3$, realized by the three vertices of an equilateral triangle. Erdős and Fishburn [2] introduce the problem of determining $g(k)$. To date only six values of $g(k)$ are known: $g(1) = 3$, $g(2) = 5$, $g(3) = 7$, $g(4) = 9$, $g(5) = 12$, and $g(6) = 13$. The first five of these were found by Erdős and Fishburn [2] and the last value was recently determined by Wei [7].

The triangular lattice is defined as follows:

$$L_\Delta = \left\{ a(1,0) + b\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) : a, b \in \mathbb{Z} \right\}.$$

Given positive integer a and non-negative integers r_1 and r_2 with $0 \leqslant r_2 \leqslant a + r_1 - 1$, let P_{a,r_1,r_2} denote a vertically symmetric contiguous and convex subset of L_Δ with $\left(\binom{r_1}{2} - \binom{r_2}{2} + a(r_1 + r_2) + r_1 - r_2 + r_1 r_2 + a\right)$

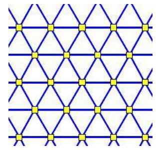

Figure 1: The triangular lattice L_Δ

points, where rows from top to bottom contain

$$a, a+1, a+2, \ldots, a+(r_1-1), a+r_1, a+r_1-1, a+r_1-2, \ldots, a+r_1-r_2$$

points, respectively. For $X \subset L_\Delta$, let $D(X)$ denote the number of distinct distances in X.

In this note, we prove two recent conjectures by Ahmed and Snevily [1].

2 Results

Conjecture 2.1 (Ahmed and Snevily[1])**.** If there exist a pair of points in P_{a,r_1,r_2} at distance d from one another, then there exist a pair of boundary points of P_{a,r_1,r_2} at a distance d from one another.

2.1 Conjecture 2.1 is true

For brevity, we shall refer to three nonparallel unit vectors in the lattice as e_1, e_2, and e_3; for instance, we might arbitrarily establish $e_1 = (1, 0)$, $e_2 = (1/2, \sqrt{3}/2)$, and $e_3 = (-1/2, \sqrt{3}/2)$.

Definition 2.1. The *boundary* of a set $S \subseteq L_\Delta$ is the set of all $x \in S$ such that at least one of the six points $x \pm e_i$ is not in S.

Definition 2.2. A nonempty set $S \subseteq L_\Delta$ is *cycle-bounded* if the boundary of S is a finite set of vertices $\{x_1, x_2, \ldots, x_n\}$ such that each $|x_{i+1} - x_i| = 1$ and $|x_1 - x_n| = 1$.

Proposition 2.1. If a finite set $S \subset L_\Delta$ is cycle-bounded, then for any $a, b \in S$, there are vertices x and y on the boundary of S such that $a - b = x - y$.

Proof. Let us start by noting that the case $a = b$ is trivial, and any $x = y$ would suffice. Henceforth, we shall assume that $a \neq b$.

Consider the sequences $a, a+e_1, a+2e_1, a+3e_1, \ldots$ and $b, b+e_1, b+2e_1, b+3e_1, \ldots$. Clearly, $a \in S$ and $b \in S$, but since S is finite, there must be some first values k and ℓ such that $a + ke_1$ and $b + \ell e_1$ are not in S.

WLOG let $k \leqslant \ell$; then we know that $a+(k-1)e_1 \in S$ and $b+(k-1)e_1 \in S$. Furthermore, since $[a+(k-1)e_1]+e_1 \notin S$, we know that $a+(k-1)e_1$ is on the boundary of S.

Now, let us color each point of the boundary red or green according to the following scheme: for a point z on the boundary of S, let z be red if $z+(b-a) \notin S$, and green if $z+(b-a) \in S$.

We may observe that the specific case of $z = a+(k-1)e_1$ is green, since $z+(b-a) = b+(k-1)e_1$, which was shown above to be in S. We may also determine that at least one point on the boundary is red, since, given that $a \neq b$, we know that $b-a \neq 0$, so it must have at least one of a positive x-component, a negative x-component, a positive y-component, or a negative y-component. In these respective cases, let us consider a vertex $z \in S$ of maximal x-coordinate, minimal x-coordinate, maximal y-coordinate, or minimal y-coordinate. These are clearly on the boundary, as can be seen by considering the coordinates of e_1, $-e_1$, e_2, and $-e_2$, respectively in the four cases. However, since $z+(b-a)$ will have greater or lesser magnitude in the appropriate coordinate, we see that $z+(b-a) \notin S$, so this point is red.

Each point on the boundary is red or green, with at least one of each, and the cycle-bounding induces an order on boundary points. Thus we may assert that there are adjacent points x_i and x_{i+1} on the boundary such that x_i is green and x_{i+1} is red. Let $x = x_i$ and $y = x_i + (b-a)$. By adjacency, we know that $x_{i+1} = x_i \pm e_j$ for some j, and by our coloring scheme, we know that $y = x_i + (b-a) \notin S$ while $y \pm e_j = x_{i+1} + (b-a) \notin S$. Thus both x and y are on the boundary, and by construction $x - y = a - b$. □

2.2 On the size of $D(P_{r+1,r,r})$

Being guided by the discussion in Ahmed and Snevily [1], we compute the number of distinct distances in $P_{r+1,r,r}$. Consider the triangular array with points from $P_{r+1,r,0}$ on the boundary or interior to the triangle, say $T_{r+1,r,0}$, with vertices

$$(0,0), (2r,0), \text{ and } (3r/2, (\sqrt{3}/2)r).$$

Let $p(i,j)$ denote the j-th point on the i-th row inside $T_{r+1,r,0}$, considering the base of $T_{r+1,r,0}$ as the 0-th row. Let $f(i,j)$ be the square of the distance from $(0,0)$ to $p(i,j)$. For example, in $T_{11,10,0}$ (Figure 2), $f(2,5) = 52$.

It can be observed that the number of points in $T_{r+1,r,0}$ is r^2+2r. Then for $0 \leqslant i \leqslant r$ and $1 \leqslant j \leqslant (2r+1)-2i$,

$$f(i,j) \quad = \quad \left(\frac{3}{2}i + j - 1\right)^2 + \left(\frac{\sqrt{3}}{2}i\right)^2.$$

The value of a node $p(i,j)$ may be repeated inside $T_{r+1,r,0}$. For example:

- In $T_{11,10,0}$, we have $f(0,8) = f(3,3) = 49$; $f(3,12) = f(7,5) = 247$.
- In $T_{16,15,0}$ (see Appendix), we have $f(4,20) = f(7,15) = f(12,6) = 637$.

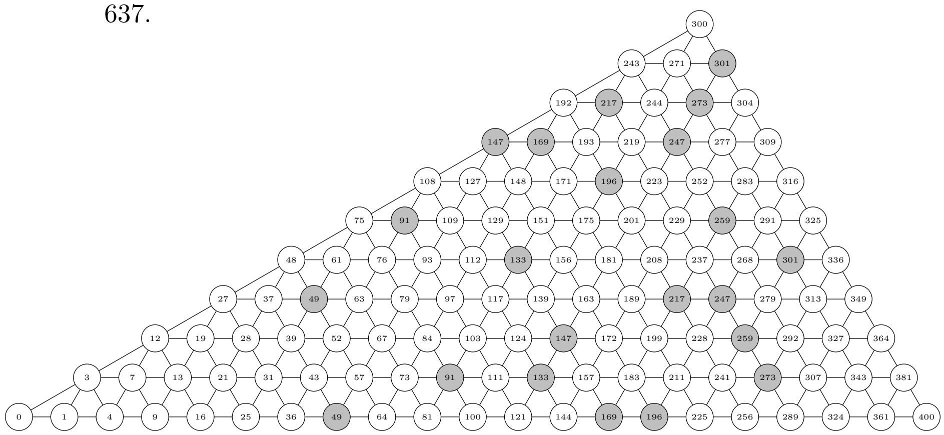

Figure 2: $T_{11,10,0}$

Observation 2.1. Inside $T_{r+1,r,0}$, we have the following:

(i) For $0 \leqslant i \leqslant r$ and $1 \leqslant j \leqslant 2r - 2i$,

$$f(i, j+1) = f(i,j) + 3i + 2j - 1 > f(i,j).$$

(ii) For $0 \leqslant i \leqslant r-1$ and $3 \leqslant j \leqslant (2r+1) - 2i$,

$$f(i+1, j-2) = f(i,j) - j + 2 < f(i,j).$$

(iii) For $0 \leqslant i \leqslant r-1$ and $2 \leqslant j \leqslant 2r - 2i$,

$$f(i+1, j-1) = f(i,j) + 3i + j > f(i,j).$$

(iv) For $0 \leqslant i \leqslant r-2$ and $4 \leqslant j \leqslant 2r - 2i$,

$$f(i+2, j-3) = f(i,j) + 3i + 3 > f(i,j).$$

Let $s(r,i,j)$ be the number of times $f(i,j)$ is repeated in $T_{r+1,r,0}$ above $p(i,j)$. Hence for $r \geqslant 1$,

$$\begin{aligned} |D(P_{r+1,r,r})| = |D(T_{r+1,r,0})| &= r^2 + 2r - \sum_{i=0}^{r} \sum_{j=1}^{2r+1-2i} v(r,i,j) \\ &= r^2 + 2r - \mathcal{X}(r) \text{ (say)}, \end{aligned}$$

where

$$v(r,i,j) = \begin{cases} 1, & \text{if } s(r,i,j) > 0; \\ 0, & \text{otherwise.} \end{cases}$$

The idea behind using $v(r,i,j)$ to indirectly compute the total number of repeated node values inside $T_{r+1,r,0}$ is as follows: suppose the node-value $f(i,j)$ appears q times in $T_{r+1,r,0}$ as

$$f(i_0,j_0), f(i_1,j_1), \ldots, f(i_{q-1},j_{q-1}),$$

with $i_0 < i_1 < \cdots < i_{q-2} < i_{q-1}$. Then

$$\begin{aligned} s(r,i_0,j_0) &= q-1 \\ s(r,i_1,j_1) &= q-2 \\ &\cdots \\ s(r,i_{q-2},j_{q-2}) &= 1 \\ s(r,i_{q-1},j_{q-1}) &= 0 \end{aligned}$$

Here, $v(r,i_0,j_0)$ indicates that $f(i,j)$ is repeated $s(r,i_0,j_0)$ times above $p(i_0,j_0)$; $v(r,i_1,j_1)$ indicates that $f(i,j)$ is repeated $s(r,i_1,j_1)$ times above $p(i_1,j_1)$; and so on. The over-count that we want to subtract from r^2+2r for the node-value $f(i,j)$ throughout $T_{r+1,r,0}$ is

$$\sum_{k=0}^{q-1} v(r,i_k,j_k) = q-1.$$

2.2.1 Computing the exact value of $s(r,i,j)$

Let $\ell(i,j)$ be the straight line passing through $p(i,j)$ making an angle $2\pi/3$ with the positive x-axis. Let $V(i,j)$ denote the vertices on $\ell(i,j)$ and inside $T_{r+1,r,0}$. Note that,

$$\ell(i,j) \equiv \ell(0,j+2i).$$

Let $d_{\min}(i,j,k)$ be the smallest square of distance of any point in $T_{r+1,r,0}$, on the line $\ell(i+k,j-k)$ from $(0,0)$. For example,

$$\begin{aligned} d_{\min}(0,8,1) &= \min\{f(i,j) : p(i,j) \in V(0+1,8-1)\} \\ &= \min\{64,57,52,49,48\} = 48. \end{aligned}$$

Observation 2.2. It can be observed that for $1 \leqslant j \leqslant 2r+1$, the number of nodes inside $T_{r+1,r,0}$ on the line $\ell(0,j)$ is $\lceil j/2 \rceil$. The node values on the line $\ell(0,j)$ are

$$f(0,j), f(1,j-2), f(2,j-4), \ldots, f(\lceil j/2 \rceil - 1, j - 2\cdot(\lceil j/2 \rceil - 1)).$$

Also for $0 \leqslant i \leqslant \lceil j/2 \rceil - 2$,

$$f(i,j) - f(i+1, j-2) = j - 2.$$

Lemma 2.1. For $0 \leqslant i \leqslant r$, $1 \leqslant j \leqslant (2r+1) - 2i$, and $1 \leqslant k \leqslant 2r - (j + 2i) + 1$,

$$d_{\min}(i,j,k) = \begin{cases} 3(i+t)^2, & \text{if } j+k = 2t+1,\ t \geqslant 1; \\ 3(i+t)^2 - 3(i+t) + 1, & \text{if } j+k = 2t,\ t \geqslant 1. \end{cases}$$

Proof. Here,

$$\ell(i+k, j-k) \equiv \ell(0, j-k+2(i+k)) \equiv \ell(0, j+2i+k)$$

and

$$f(0, j+2i+k) = (j+2i+k-1)^2 .$$

Note that for $i = 0$, we have $j + k \leqslant 2r+1$. Now, we have the following two cases:

(*i*) $j+k = 2t+1$ for some $1 \leqslant t \leqslant r$: Since $j+k$ is odd, we have $j+2i+k$ odd. So, number of nodes in $T_{r+1,r,0}$ on the line $\ell(0, j+2i+k)$ is

$$\lceil (j+2i+k)/2 \rceil = \lceil (2t+1+2i)/2 \rceil = i+t+1.$$

Therefore, using Observation 2.2,

$$\begin{aligned} d_{\min}(i,j,k) &= f(0, j+2i+k) - \sum_{q=0}^{i+t-1} (f(q, (j+2i+k) - 2q) \\ &\quad - f(q+1, (j+2i+k) - 2(q+1))) \\ &= (2i+j+k-1)^2 - \sum_{q=0}^{i+t-1} ((j+2i+k) - 2q - 2) \\ &= (2i+j+k-1)^2 - (j+2i+k-2)(i+t) - 2\sum_{q=0}^{i+t-1} q \\ &= (2i+2t+1-1)^2 - (i+t)^2 = 3(i+t)^2. \end{aligned}$$

(*ii*) $j+k = 2t$ for some $1 \leqslant t \leqslant r$: Here, the number of nodes in $T_{r+1,r,0}$ on the line $\ell(0, j+2i+k)$ is $\lceil (j+2i+k)/2 \rceil = \lceil (2t+2i)/2 \rceil = i+t$.

Hence,

$$\begin{aligned}
d_{\min}(i,j,k) &= f(0, j+2i+k) - \sum_{q=0}^{i+t-2} (f(q, (j+2i+k) - 2q) \\
&\quad - f(q+1, (j+2i+k) - 2(q+1))) \\
&= (2i+j+k-1)^2 - \sum_{q=0}^{i+t-2} ((j+2i+k) - 2q - 2) \\
&= (2i+2t-1)^2 - (i+t)(i+t-1) \\
&= 3(i+t)^2 - 3(i+t) + 1.
\end{aligned}$$

□

For $0 \leqslant i \leqslant r$, $1 \leqslant j \leqslant (2r+1) - 2i$, and $1 \leqslant k \leqslant 2r - (j+2i) + 1$, define

$$c(i,j,k) = \begin{cases} 0, & \text{if } d_{\min}(i,j,k) > f(i,j); \\ b(i,j,k), & \text{otherwise.} \end{cases}$$

where

$$b(i,j,k) = \begin{cases} 1, & \text{if } j+k = 2t+1,\ f(i,j) - d_{\min}(i,j,k) = y^2 \\ & \text{with some } 0 \leqslant y \leqslant t; \\ 0, & \text{if } j+k = 2t+1,\ f(i,j) - d_{\min}(i,j,k) \neq y^2 \\ & \text{for any } 0 \leqslant y \leqslant t; \\ 1, & \text{if } j+k = 2t,\ f(i,j) - d_{\min}(i,j,k) = y(y+1) \\ & \text{with some } 0 \leqslant y < t; \\ 0, & \text{if } j+k = 2t,\ f(i,j) - d_{\min}(i,j,k) \neq y(y+1) \\ & \text{for any } 0 \leqslant y < t. \end{cases}$$

The function $c(i,j,k)$ indicates if $f(i,j)$ is a node value on the line $\ell(0, j+2i+k)$. Clearly, $c(i,j,k) = 0$ if $d_{\min}(i,j,k) > f(i,j)$. Otherwise, $f(i,j)$ may or may not be a node value on the line $\ell(0, j+2i+k)$. In either case, the line $\ell(0, j+2i+k)$ contains node values greater than $f(i,j)$ in $T_{r+1,r,0}$. Consider the following cases:

(*i*) $j+k = 2t+1$ for $1 \leqslant t \leqslant r$: There are $i+t+1$ nodes on the line $\ell(0, j+2i+k)$. It can be observed that $d_{\min}(i,j,k) = f(i+t, 1)$. The y-th ($0 \leqslant y \leqslant t$) node-value from $d_{\min}(i,j,k)$ downwords along the line $\ell(0, j+2i+k)$ is $f(i+t-y, 1+2y)$. It can be verified that

$$f(i+t-y, 1+2y) - f(i+t, 1) = y^2.$$

So, $f(i,j) - d_{\min}(i,j,k) = y^2$ for some y with $0 \leqslant y \leqslant t$ indicates the existence of the node-value $f(i,j)$ on the line $\ell(0, j+2i+k)$.

(*ii*) $j{+}k = 2t$ for $1 \leqslant t \leqslant r$: There are $i{+}t$ nodes on the line $\ell(0, j{+}2i{+}k)$. It can be observed that $d_{\min}(i,j,k) = f(i{+}t{-}1, 2)$. The y-th ($0 \leqslant y < t$) node-value from $d_{\min}(i,j,k)$ downwords along the line $\ell(0, j{+}2i{+}k)$ is $f(i+t-y-1, 2+2y)$. It can be verified that

$$f(i+t-y-1, 2+2y) - f(i+t-1, 2) = y(y+1).$$

So, $f(i,j) - d_{\min}(i,j,k) = y(y{+}1)$ for some y with $0 \leqslant y < t$ indicates the existence of the node-value $f(i,j)$ on the line $\ell(0, j+2i+k)$.

Then

$$s(r,i,j) = \sum_{k=1}^{2r-(j+2i)+1} c(i,j,k).$$

2.2.2 Asymptotic bound for $\mathcal{X}(r)$

Recall that

$$\mathcal{X}(r) = \sum_{i=0}^{r} \sum_{j=1}^{2r+1-2i} v(r,i,j).$$

We get the following experimental values:

r	5	6	7	8	9	10	11	12	13	14	15
$\mathcal{X}(r)$	1	2	4	6	9	11	14	18	23	28	33
r	17	18	19	20	21	22	23	24	25	26	27
$\mathcal{X}(r)$	46	52	60	68	78	88	98	108	118	130	144
r	29	30	31	32	33	34	35	36	37	38	39
$\mathcal{X}(r)$	168	184	201	217	230	248	264	280	304	323	342
r	50	60	70	80	90	100	120	150	200		
$\mathcal{X}(r)$	608	924	1312	1775	2310	2913	4363	7124	13320		

$\chi(r)$ seems to grow far faster than early indications would suggest. Among small values of r, multiple points within the triangle at equal distance from $(0,0)$ seem quite rare, but experimental data suggests that duplications may well be of quadratic frequency. This is in fact the case, as can easily be determined from a reconsideration of a known asymptotic result. The ability of a natural number to be the square of a distance in a hexagonal lattice was characterized by Marshall [6] as equivalent to possessing no prime factor congruent to 2 modulo 3 an odd number of times; these numbers are called *Löschian* after the economist Augustus Lösch [5]. Except for the modulus and residue in that congruence, this criterion is identical to Fermat's characterization of numbers which are sums of two squares.

Fermat's result is integral to the proof of Landau [3], presented in an elementary form in [4], that the number of sums of two squares which are less than n is

$$K\frac{n}{\sqrt{\log n}} + O\left(\frac{n}{\log^{3/4} n}\right)$$

for constant

$$K = \sqrt{\frac{1}{2} \prod_{p \equiv 3 \pmod 4} \frac{1}{1 - p^{-2}}}.$$

Landau's line of argument is easily adaptable to a different modulus and residue, from which the density of the Löschian numbers may be shown to be

$$L \frac{n}{\sqrt{\log n}} + O\left(\frac{n}{\log^{3/4} n}\right)$$

for constant

$$L = \sqrt{\frac{1}{2\sqrt{3}} \prod_{p \equiv 2 \pmod 3} \frac{1}{1 - p^{-2}}}.$$

Since $P_{r+1,r,r}$ lies entirely within a circle of radius $2r$ and entirely contains a circle of radius $\sqrt{3}r$, we may say with certainty that $|D(P_{r+1,r,r})|$ lies between the number of Loeschian numbers less than or equal to $3r^2$ and the number of Loeschian numbers less than or equal to $4r^2$. Thus,

$$L \frac{3r^2}{\sqrt{\log(3r^2)}} \leqslant |D(P_{r+1,r,r})| \leqslant L \frac{4r^2}{\sqrt{\log(4r^2)}} + O\left(\frac{r^2}{\log^{3/4} r}\right)$$

Since $\chi(r) = r^2 + 2r - D(|P_{r+1,r,r}|)$, it is thus abundantly clear that $\chi(r) = r^2 + 2r - O\left(\frac{r^2}{\sqrt{\log r}}\right)$.

References

[1] Tanbir Ahmed and Hunter Snevily, Sparse Distance Sets in the Triangular Lattice, *Elec. J. of Comb.*, **20(4)** (2013), #P33.

[2] P. Erdős and P. Fishburn, Maximum planar sets that determine k distances, *Discrete Math.*, **160** (1996), 115–125.

[3] Edmund Landau, Uber die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindeszahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate, *Archiv der Mathematik und Physik*, **13** (1908), 305–312.

[4] William Judson LeVeque, *Topics in number theory*, Addison-Wesley Pub. Co., 1956.

[5] A. Lösch, *The Economics of Location*, Yale University Press, 1954.

[6] John U. Marshall, The Löschian Numbers As a Problem in Number Theory, *Geographical Analysis*, **7(4)** (1975), 421–426.

[7] X. Wei, A proof of Erdős-Fishburn's conjecture for $g(6) = 13$, *Elec. J. of Comb.*, **19(4)** (2012), #P38.

A $T_{16,15,0}$