SERIE B — INFORMATIK

On point Sets with Many Unit Distances in Few Directions

Peter Braß*

B 97-04
June 1997

Abstract

We study the problem of the maximum number of unit distances among n points in the plane under the additional restriction that we count only those unit distances that occur in a fixed set of k directions, taking the maximum over all sets of n points and all sets of k directions. We prove that for fixed k and sufficiently large $n > n_0(k)$ the extremal sets are essentially sections of lattices, bounded by edges parallel to the k directions and of equal length.

* Institut für Informatik, Freie Universität Berlin, Takustraße 9, D-14195 Berlin, Germany
email: brass@inf.fu-berlin.de

On Point Sets with Many Unit Distances in Few Directions

Peter Brass

FU Berlin, Germany

Dedicated to the Memory of Paul Erdős

Abstract: We study the problem of the maximum number of unit distances among n points in the plane under the additional restriction that we count only those unit distances that occur in a fixed set of k directions, taking the maximum over all sets of n points and all sets of k directions. We prove that for fixed k and sufficiently large $n > n_0(k)$ the extremal sets are essentially sections of lattices, bounded by edges parallel to the k directions and of equal length.

1. Introduction

The problem of the maximum number of unit distances is perhaps the simplest-sounding of Erdős's many combinatorial geometry problems; nonetheless it turned out to be a very difficult problem which is still far from its solution. The question 'How often can the same distance appear among n points in the plane' was first studied by Erdős in his 1946 paper [9], where he constructed an upper bound of order $O(n^{\frac{3}{2}})$ and a lower bound of order $\Omega\left(ne^{c\frac{\log n}{\log\log n}}\right)$, the last by taking a square section of a properly scaled integer lattice. The upper bound was then reduced in several steps ([1], [12]) to $O\left(n^{\frac{4}{3}}\right)$ [16]. Since the first proof was very complicated, alternative proofs were sought ([5] by 'random resampling', [13] by use of the VC-dimension) until 1996 Székely found a very elegant proof using a crossing-number argument [17].

There has been no change in the lower bound: although triangular lattice sections seem to give slightly better numbers than square lattice sections, they have the same asymptotics, and lattice sections are still the best known construction. Erdős repeatedly offered 250\$ for an upper bound which is $O(n^{1+\epsilon})$ for all $\epsilon > 0$ and 500\$ for an upper bound $O\left(ne^{c\frac{\log n}{\log\log n}}\right)$. Exact values and extremal sets for up to 14 points (Figure 1) were determined by Schade ([15], see also [2]). Also many related problems were discussed, e.g. the maximum number of unit distances in other metric spaces, higher dimensions, convex or general position, of smallest or largest distances etc. (see [3] for further references).

The original problem, however, seems to have exhausted the available methods with the $O(n^{\frac{4}{3}})$-bound. All known proofs of this upper bound used as main geometric information the fact that any two unit circles intersect in at most two points; and this information is used best possible in that bound, since there are strictly convex norms on $\mathbb{R}^2$ (which have the same intersection pattern) for which $cn^{\frac{4}{3}}$ unit distances among n points are possible [3]. Attempts to enumerate the small excluded substructures of unit distance graphs ([8], [14]) did not yield any other excluded substructure that could be used to reduce this upper bound. Also the Schade sets do not support the implicit belief that lattice sections are extremal sets, since the larger sets are not subsets of any lattice [3], although rational dimension is still small (they are generated by four unit vectors).

In the following all points will be in the plane, and we will distinguish pointsets only up to isometry.

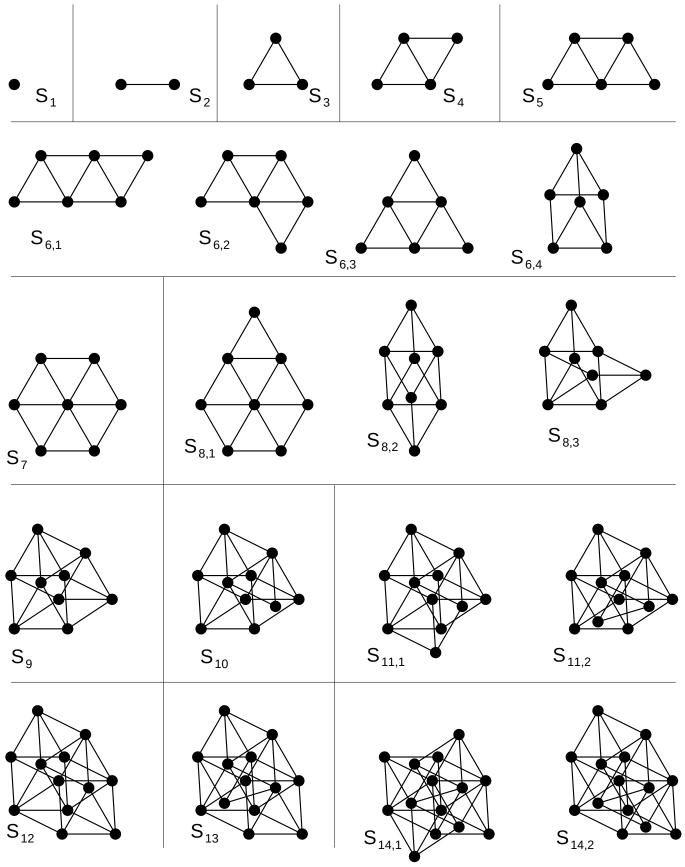

Figure 1

2. The Result

In the following we try to avoid these difficulties by proving structural results for the extremal sets. Since a lattice subset structure in itself is rather useless (for it may be a very thin subset), we aim for a lattice section structure. We call a set $S \subset \Gamma$ a section of Γ if there is a convex set K such that all points of Γ that are in the interior of K belong to S, and all points of Γ that are in the extrerior of K do not belong to S. We make no assumption on the lattice points on the boundary of S; this allows incremental construction of sections for all numbers of points in S.

To obtain results, we introduce additional structure on the pointsets with many unit distances by counting only those unit distances that occur in a few fixed directions. Let $S \subset \mathbb{R}^2$ be a set of n points and $T \subset \mathbb{R}^2$ a set of k unit vectors with $T \cap (-T) = \emptyset$ (we count each direction only once), and let

$$f(S,T) := \Big|\{(x,y) \in S \times S \mid x - y \in T\}\Big|$$

be the number of unit distances in S which occur in directions from T. We study now the function $f(n,k)$ defined as the maximum of $f(S,T)$ over all such sets S of n points and T

of k unit vectors. Trivially $f(n,1) = n-1$. The unit distance graphs with unit distances in two directions are subgraphs of the union of disjoint copies of the square lattice graph; the maximum edge-number of a subgraph of the square-lattice graph was determined by Harary and Harborth [10], so we have $f(n,2) = \lfloor 2n - 2\sqrt{n} \rfloor$. The maximum number of edges in the triangular-lattice graph follows from a disk-packing theorem by Harborth [11]; this suggests that $f(n,3) = \left\lfloor 3n - \sqrt{12n-3} \right\rfloor$. We prove

Theorem : For each fixed k we have $f(n,k) = kn - \Theta\left(\sqrt{n}\right)$, and for $k \geq 3$ there is a finite number of nonisometric lattices $\Gamma_{k,1}, \ldots, \Gamma_{k,l_k}$ such that for each sufficiently large $n > n_0(k)$ each extremal set pair $S_{n,k}, T_{n,k}$ is a pair of subsets of one of the $\Gamma_{k,i}$.
There is always an extremal set $S_{n,k}$ that is, with the exception of at most $O(\sqrt{n})$ points, a section of that $\Gamma_{k,i}$ which is bounded by edges parallel to the vectors from $T_{n,k}$.

So for fixed number of directions and large number of points the extremal sets have a lattice section structure, they are 'equilateral sections' (see Lemma 2) like the set in Figure 2. If the numbers $n_0(k)$ were small, this would imply the Erdős conjecture, since each extremal set for the unrestricted maximum number of unit distances must be an extremal set for some number of directions k; but the Schade sets show that we cannot do without the lower bounds $n_0(k)$.

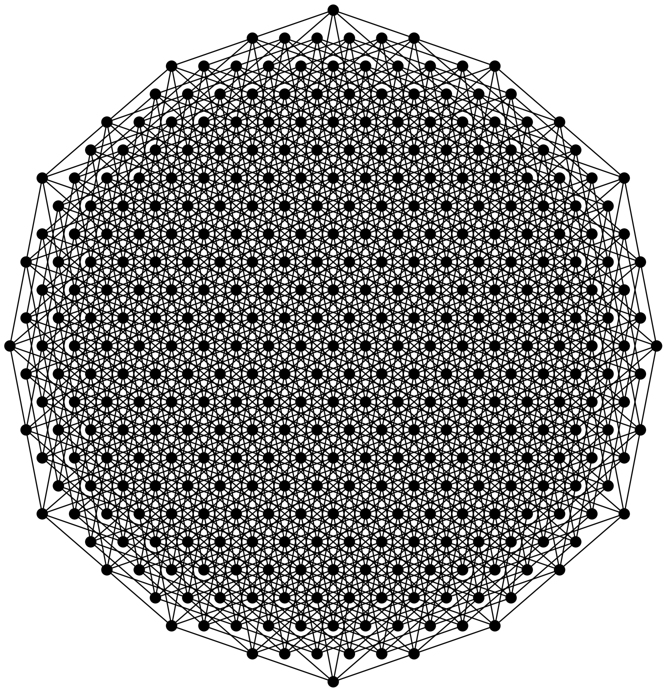

Figure 2

3. The Proof

In the following we denote for any pointset X the unit distance graph of $X \subset \mathbb{R}^2$ by $G(X)$, and if additionally a set $Y \subset \mathbb{R}^2$ of vectors is given, the graph of vertex pairs from X with a difference in Y by $G_Y(X)$. The edgenumber of graph G is written as $e(G)$.

The lower bound $f(n,k) \geq kn - O(\sqrt{n})$ is simple: Just select some r such that the unit-distance graph $G(r\mathbf{Z}^2)$ is regular of degree at least $2k$, e.g. $r = 5^{-\frac{1}{2}\lceil\frac{k}{4}\rceil}$. Then we can select a set of k unitvectors $T \subset r\mathbf{Z}^2$ (with $T \cap (-T) = \emptyset$) and a set of n points $S := r\Big(\mathbf{Z}^2 \cap [1, \sqrt{n}]^2\Big)$. In this set S each point with distance at least one to the boundary of the section (each point in $[r+1, r\sqrt{n}-1]^2$) has a neighbour in S in each direction of T, so each of these points has degree $2k$ in $G_T(S)$. Therefore $f(n,k) \geq e\Big(G_T(S)\Big) > k\Big(n - \frac{4}{r}\sqrt{n}\Big)$.

The theorem is now obtained from the following two lemmas.

Lemma 1: For each $k \geq 3$ there is a finite number of nonisometric lattices $\Gamma_{k,1}, \ldots, \Gamma_{k,l_k}$ such that for all $n > n_0(k)$ each extremal pair $(S_{n,k}, T_{n,k})$ is isometric to a pair of subsets of one of the $\Gamma_{k,i}$.

Lemma 2: Let $\Gamma \subset \mathbb{R}^2$ be a lattice and $T \subset \Gamma \setminus \{0\}$ a set of vectors such that the generated graph $G_T(\Gamma)$ is connected, and $T \cup (-T)$ does not contain three collinear vectors. Let $k := \frac{1}{2}|T \cup (-T)|$.

Let $f(n)$ denote the maximum number of edges of a subgraph of $G_T(\Gamma)$ with n vertices, and $h(n)$ denote the maximum edge-number in an n-vertex subgraph that is generated by a section of Γ. Then for $n \geq n_0(T,\Gamma)$ the extremal sections for $h(n)$ are sections bounded by edges parallel to the graph-edges which are of relative (graph-) lengths that differ by at most one, and the extremal subsets for $f(n)$ differ from the extremal sections by at most $O(\sqrt{n})$ points. Also $f(n) = h(n) + O(1) = kn - c_{T,\Gamma}\sqrt{n} + \Theta(1)$.

The assumption of no collinear generating vectors is necessary for the 'almost equilateral structure' of the extremal sections, as can be seen by the graph generated on $\mathbf{Z}^2$ by the vectors $\{(1,0),(0,1),(0,2)\}$. We conjecture that for $n \geq n_0(T,\Gamma)$ the extremal subsets are sections, so $f(n) = h(n)$.

Proof of Lemma 1: Let (S,T) denote an extremal pair for $f(n,k)$. By translation invariance we may assume $0 \in S$. If $G_T(S)$ were not connected, we could move one of the connected components against the others until a new unit vector of T occurs; but then the set was not extremal. So for extremal S the graph $G_T(S)$ is connected. Therefore each point of S can be reached from the point 0 by following T-edges; thus each point of S is an integer linear combination of T-vectors: $S \subset \langle T\rangle_Z$.

Let $T = \{t_1, \ldots, t_k\}$, so $G_T(S)$ has maximum degree $2k$. For each decomposition $T = T_1 \cup \cdots \cup T_l$ in disjoint subsets we have

$$kn - e\Big(G_T(S)\Big) = \sum_{i=1}^{l}\Big(|T_i|n - e\Big(G_{T_i}(S)\Big)\Big),$$

and each summand is positive, $e\Big(G_{T_i}(S)\Big) \leq f\Big(|S|, |T_i|\Big) < |T_i|\,|S|$. Since each t_i is a unit vector, and $T \cap (-T) = \emptyset$, any two distinct $t_i \neq t_j$ are rationally independent. Therefore each graph $G_{\{t_i,t_j\}}(S)$ is an $n = |S|$-vertex subgraph of the classical square lattice graph. $G(\mathbf{Z}^2) = P_\infty \times P_\infty$. By the theorem of Harary and Harborth [10] we have $e\Big(G_{\{t_i,t_j\}}(S)\Big) \leq \lfloor 2n - 2\sqrt{n}\rfloor$. Taking the sum over all 2-element subsets of T we find

$$e\Big(G_T(S)\Big) \leq kn - k\sqrt{n}.$$

This proves $f(n,k) = kn - \Theta\left(\sqrt{n}\right)$.

Suppose now that T contains three rationally independent unit vectors t_1, t_2, t_3. Then the graph $G_{\{t_1,t_2,t_3\}}(S)$ is isomorphic to a subgraph of the unit distance graph $G(\mathbf{Z}^3)$ of the three-dimensional integer lattice. For we can assign each point of S in each connected component of $G_{\{t_1,t_2,t_3\}}(S)$ integer coordinates relative to the basis t_1, t_2, t_3 and an arbitrary starting point of that component, with the edges corresponding to the point pairs with only one coordinate differing by exactly one. But an n-vertex subgraph of the three-dimensional lattice graph $G(\mathbf{Z}^3)$ contains less than $3n - \left(\frac{9}{2}\pi\right)^{\frac{1}{3}} n^{\frac{2}{3}}$ edges (the exact maximum number is not known; probably cubic sections are optimal. An upper bound on the number e of unit distances in a set $X \subset \mathbf{Z}^3$ of n points one obtains easily by considering the Minkowski sum $Y := X + [\frac{-1}{2}, \frac{1}{2}]^3$. If A is the surface area of Y, then $3n - \frac{1}{2}A \geq e$. The volume of Y is at least n, so by the isoperimetric theorem we have $A \geq \left(36\pi n^2\right)^{\frac{1}{3}}$, and thus $e \leq 3n - \left(\frac{9}{2}\pi\right)^{\frac{1}{3}} n^{\frac{2}{3}}$.). So if for an extremal pair (S,T) the set T contains three rationally independent vectors, then $f(n,k) \leq kn - \left(\frac{9}{2}\pi\right)^{\frac{1}{3}} n^{\frac{2}{3}}$; this is for $n > n_0(k)$ a contradiction to our lower bound $f(n,k) > kn - c_k\sqrt{n}$. Therefore the set of integer combinations $\langle T\rangle_Z$ is for fixed k and sufficiently large n always a lattice (rational dimension two), and S is a subset of this lattice.

Next we will show that for each fixed $k \geq 3$ there are only finitely many nonisometric lattices that can occur as the underlying lattice $\Gamma := \langle T\rangle_Z$ of the extremal sets (S,T) (which we again assume to be an extremal pair for $f(n,k)$). Γ is the thinnest lattice that contains S. Let now $t_a, t_b \in T$ be two distinct directions from T, then $2n - 2\sqrt{n} \geq e\left(G_{\{t_a,t_b\}}(S)\right) \geq 2n - c_k\sqrt{n}$, since the defect in two directions is smaller than the total defect $c_k\sqrt{n}$. The vectors t_a, t_b generate a sublattice $\langle t_a, t_b\rangle_Z \subset \Gamma$, and $\Gamma/\langle t_a, t_b\rangle_Z$ consists of a finite number $r = r(t_a, t_b)$ of cosets $\Gamma_1, \ldots, \Gamma_r$, which are translates of the sublattice $\langle t_a, t_b\rangle_Z$. The remaining unit vectors t_i $(i \neq a, b)$ of T operate on these cosets, since they are lattice vectors in Γ; so for each Γ_j, t_i the sum set $\Gamma_j + t_i$ is another of those finitely many cosets. But at least $\left|S\cap\Gamma_j\right| - \left|S\cap(\Gamma_j + t_i)\right|$ of the points of $S \cap \Gamma_j$ have no neighbour in S in the direction t_i. Taking the sum over all cosets Γ_j and all directions t_i we find

$$\sum_{j=1}^{r}\sum_{i=1}^{k}\left|\left|S\cap\Gamma_j\right| - \left|S\cap(\Gamma_j+t_i)\right|\right| \leq kn - f(n,k) \leq c_k\sqrt{n}.$$

Since T generates Γ, the graph generated by the t_i on the cosets is connected; so we get from the previous inequality especially

$$\max_{i=1}^{r}\left|S\cap\Gamma_i\right| - \min_{i=1}^{r}\left|S\cap\Gamma_i\right| \leq c_k\sqrt{n}.$$

Let now $n_i := \left|S\cap\Gamma_i\right|$. Then the graph $G_{\{t_a,t_b\}}\left(S\cap\Gamma_i\right)$ contains between the n_i vertices at most $\left\lfloor 2n_i - 2\sqrt{n_i}\right\rfloor$ edges. Since $G_{\{t_a,t_b\}}(S)$ contains at least $2n - c_k\sqrt{n}$ edges, we have

$$\sum_{i=1}^{r}\sqrt{n_i} \leq \frac{c_k}{2}\sqrt{n}. \qquad (*)$$

We now look for the minimum of $\sum_{i=1}^{r}\sqrt{n_i}$ under the restrictions $\sum_{i=1}^{r} n_i = n$, $n_i \geq 0$ for $i = 1, \ldots, r$ and $\max_{i=1}^{r} n_i - \min_{i=1}^{r} n_i \leq c_k\sqrt{n}$. Since $\sqrt{x}$ is a concave function, for the extremal choice of the n_i each variable (with at most one exception) will meet a restriction of the admissible set. So under the n_i there are only two distinct values,

$$\max_{i=1}^{r} n_i =: \frac{n}{r} + \alpha c_k \sqrt{n} \quad \text{and} \quad \min_{i=1}^{r} n_i =: \frac{n}{r} + (\alpha - 1) c_k \sqrt{n},$$

where the last has to be nonnegative.

If $\min_{i=1}^{r} n_i \geq \frac{n}{2r}$, then we have

$$\sum_{i=1}^{r} \sqrt{n_i} > r \sqrt{\min_{i=1}^{r} n_i} \geq r \sqrt{\frac{n}{2r}} = \sqrt{\frac{r}{2}} \sqrt{n}.$$

By $(*)$ we find that in this case r is bounded by a constant for fixed k and sufficiently large n.

If $0 \leq \min_{i=1}^{r} n_i < \frac{n}{2r}$, that is $0 \leq \frac{n}{r} + (\alpha - 1) c_k \sqrt{n} < \frac{n}{2r}$, we obtain $\max_{i=1}^{r} n_i = \min_{i=1}^{r} n_i + c_k \sqrt{n} \geq c_k \sqrt{n}$, and $(1 - \alpha) \geq \frac{\sqrt{n}}{2 c_k r}$. Since $\sum_{i=1}^{r} n_i = n$ holds, the smaller value of the n_i must occur αr times and the larger value $(1 - \alpha) r$ times. Therefore we have

$$\begin{aligned} \sum_{i=1}^{r} \sqrt{n_i} &= (1 - \alpha) r \sqrt{\max_{i=1}^{r} n_i} + \alpha r \sqrt{\min_{i=1}^{r} n_i} \\ &\geq (1 - \alpha) r \sqrt{c_k \sqrt{n}} \geq \frac{\sqrt{n}}{2 c_k r} r \sqrt{c_k \sqrt{n}} \\ &\geq \frac{1}{2 \sqrt{c_k}} n^{\frac{3}{4}}. \end{aligned}$$

By $(*)$ we find that this case cannot occur for large n.

So for fixed k and sufficiently large n there are only finitely many possible values for $r(\Gamma, T) = \max r(t_a, t_b)$. To show now that for bounded $r(\Gamma, T)$ and $k \geq 3$ there are only finitely many nonisometric lattices Γ possible, we look at the sublattices of Γ generated by three vectors $t_{i_1}, t_{i_2}, t_{i_3}$. If all these sublattices are known (up to isometry), then Γ is determined (up to isometry), since Γ is generated by T. Since $t_{i_1}, t_{i_2}, t_{i_3}$ are vectors of a lattice, they are rationally dependent, so there are relatively prime integers a_1, a_2, a_3 such that $a_1 t_{i_1} + a_2 t_{i_2} + a_3 t_{i_3} = 0$. These integers a_i are nonzero, since $T \cap (-T) = \emptyset$, $t_i \neq \pm t_j$ for $i \neq j$. By the coefficients a_1, a_2, a_3 and the fact that all t_i are unit vectors, the lattice $\langle t_{i_1}, t_{i_2}, t_{i_3} \rangle_Z$ is determined up to isometry, since $\|t_{i_1}\| = 1$, $\|t_{i_2}\| = 1$ and $\cos(\sphericalangle t_{i_1}, t_{i_2}) = \left(t_{i_1}, t_{i_2}\right) = \frac{a_3^2 - a_2^2 - a_1^2}{2 a_1 a_2}$. We now take the sublattice generated by two vectors t_{i_1} and t_{i_2} of the lattice generated by all three vectors $\langle t_{i_1}, t_{i_2}, t_{i_3} \rangle_Z$, which itself is a sublattice of Γ. The number of cosets of $\langle t_{i_1}, t_{i_2} \rangle_Z$ in Γ is at least as large as the number of cosets of $\langle t_{i_1}, t_{i_2} \rangle_Z$ in $\langle t_{i_1}, t_{i_2}, t_{i_3} \rangle_Z$; but that is $|a_3|$. Therefore the coefficients a_1, a_2, a_3 are bounded in absolute value by r. So for each sublattice of Γ generated by three unit vectors there are at most $(2r)^3$ nonisometric possibilities. So for fixed k and sufficiently large n there are only finitely many nonisometric lattices Γ possible. This proves Lemma 1.

Proof of Lemma 2: Let S_n and X_n denote the extremal subset and section of n points, respectively. Each point of an extremal set (X_{n+1} or S_{n+1}) that is a vertex of the convex hull of that set has degree at most k in the graph, since all its $T \cup (-T)$-neighbours are on one side of a supporting line through that point. Removing that point gives $f(n+1) \leq f(n) + k$ and $h(n+1) \leq h(n) + k$ for all n. We are therefore most interested in the numbers n for which f or h grow by at most $k-1$.

In the first step we show that in the extremal sets (X_n or S_n) most points have a big neighbourhood in which no point of Γ is missing in the extremal set.

A square section of Γ gives a lower bound $f(n) \geq h(n) \geq kn - O(\sqrt{n})$. Therefore in S_n as well as in X_n all points with $O(\sqrt{n})$ exceptions have full degree $2k$, i.e. if p belongs to the extremal set, then also $\{p \pm t_i \mid i = 1, \ldots, k\}$. For each fixed distance d_G in the graph, the $O(\sqrt{n})$ points that do not have full degree have only $O(\sqrt{n})$ neighbours with graph-distance (in $G_T(\Gamma)$) at most d_G; so for each fixed r_G all but $O(\sqrt{n})$ points in the extremal sets are centers of balls of radius r_G (in the graph-metric of $G_T(\Gamma)$) that completely belong to the extremal sets. For each euclidean ball with euclidean radius r_e around a point of Γ there is a number r_G such that the ball of radius r_G in the graph metric contains the euclidean ball of radius r_e. Therefore the same statement also holds for the euclidean metric: for each fixed euclidean distance d_e all but $O(\sqrt{n})$ points of the extremal set have distance at least d_e to the next lattice point outside the extremal set.

In the next step we show that in each extremal set most points are contained in sections that are bounded by edges parallel to the vectors of T.

Let P denote the open 0-symmetric convex $2k$-gon, whose edgevectors are the vectors of $T \cup (-T)$ and let Z be the union of all those translates of P that do not contain a lattice point from outside the extremal set ($\Gamma \setminus X_n$ or $\Gamma \setminus S_n$, respectively). Since P has finite euclidean diameter, all but $O\left(\sqrt{n}\right)$ points of the extremal set are centers of a translate of P which is completely inside the extremal set, so Z contains all but $O\left(\sqrt{n}\right)$ points of the extremal set. Z consists of finitely many connected components, each of which is an open polygon, bounded by edges parallel to the directions of T. We denote these connected components by $Z_1, \ldots, Z_l$. In the following we denote as the relative length of a segment with direction $t \in T$ the ratio of its length to that of t.

We claim that for those n for which the function associated with the extremal set (h or f) grows by less than k, these connected components must indeed be convex polygons. For by the construction of P we have the property that if q is a point outside P and $t \in T \cup (-T)$ is a vector such that the ray with direction t starting in q intersects the interior of P, then at least one of the points $q+t, q+2t, q+3t, \ldots$ is in the interior of P. This property carries over to any union of translates of P, so especially to each connected component of $Z + P$.

Suppose now that one of the connected components Z_i is not convex, and increase the extremal set by all those points inside the smallest open convex polygon that contains Z_i and is bounded by edges parallel to the directions of T. Let ν be the number of added vertices and η the number of added edges; we claim $\eta \geq k\nu$, so $f(n+\nu) \geq f(n) + k\nu$ which contradicts the assumption on n. To count the number of additional edges we note that each edge $\{q, q+t_i\}$ belongs to exactly one infinite sequence $(q + at_i)_{a=-\infty}^{\infty}$ (a one-dimensional sublattice), and each such sequence that intersects the enlarged polygon already intersects the original polygon Z_i. So if such a sequence contains new points, it already contained at least one old point of the extremal set, but the points of the enlarged set form an interval in that sequence. Then in that sequence there are at least as many new edges as new points. But each new point belongs to exactly k such sequences, and each new edge to exactly one, so $\eta \geq k\nu$, which proves the claim.

Now we investigate the structure of the boundary of such a section component;

Since each Z_i is convex and the union of translates of P, each edge of Z_i (which is direction of some vector $t \in T$) contains at least one edge of the translate of P touching that edge of Z_i, so it is at least as long as that vector t. We call this edge of Z_i trivial if it is exactly as long as t, and nontrivial if it is longer. A nontrivial edge $\overline{a(a+\sigma t)}$ must contain lattice points p from the complement of the extremal set in each part of relative length greater than one, since there are points that block movement of P across that edge in each part of the edge. Since p is a lattice point and t a lattice vector, there is at least one one-dimensional sublattice $\dots p-t, p, p+t, \dots$ of which a segment $\Sigma := \{p + zt \mid z \in \mathbf{Z}\} \cap \overline{a(a+\sigma t)}$ of $\lfloor \sigma \rfloor$ or $\lfloor \sigma \rfloor + 1$ points is contained in the edge (if t is not a primitive lattice-vector, there will be several such sublattices). Σ defines a path in $G_T(\Gamma)$, and each of the points of Σ (with the possible exception of the end point a and $a + \sigma t$ of the segment) has already $k-1$ neighbours in the extremal set. So if there is a point of Σ already belonging to the extremal set, and it has a neighbour in Σ (that is not one of the endpoints of the segment), then this neighbour could be added to the extremal set giving k new edges. So for those n in which the associated function (f or h) grows only by $k-1$, each such Σ must either completely belong to the extremal set, or completely belong to its complement; and each nontrivial edge of a Z_i will contain at least one such Σ that completely belongs to the complement of the extremal set (since there are missing points on the edge that stop movement across it).

We also note that the existence of nontrivial edges implies the possibility to add a point of degree at least $k-1$; and in an extremal set there have to be nontrivial edges. For a connected component with only trivial edges contains only $O(1)$ points, among them several that are not of full degree $2k$ in the graph of the extremal set. But there are at most $O(\sqrt{n})$ such points, so not all n points of the extremal set can be contained in connected components with only trivial edges. So for all $n \geq n_0(T, \Gamma)$ we have $k \geq f(n+1) - f(n) \geq k-1$ and $k \geq h(n+1) - h(n) \geq k-1$.

With this information we can now determine the structure of the extremal sections X_n and thus prove the first statement of the lemma. Since the extremal sections are generated by convex sets, they can contain only one connected component Z. If there are points of X_n in the exterior of this component, then any edge of the component that separates these points from Z must have relative length at most two, for otherwise the convexity of the generating set would force a segment of relative length one of the edge to contain only points of X_n, contradicting the existence of points of $\Gamma \setminus X_n$ in every edge-segment of relative length one. But there are only $O(1)$ points beyond such a short edge possible, since they must lie in the triangle bounded by the lines through the previous and following edge. So we can remove these l_1 points beyond the short edge (losing at most k edges per points, since we can remove them in the order in which they become available as vertices of the convex hull) as well as the l_2 points on the short edge (losing at most $kl_2 - 1$ edges, since they form a union of paths, each point with k_1 neighbours on the other side of the short edge), with $l_1 + l_2 = O(1)$. But since Z contains all but $O(\sqrt{n})$ points of X_n, there must be a side of length $\Omega(\sqrt{n})$, along we can select a missing one-dimensional sublattice Σ as described above. This Σ determines a path of length $\Omega(\sqrt{n})$ in $G_T(\Gamma \setminus X_n)$, of which we can add a subpath of length $l_1 + l_2$ to the set X_n (replacing the points on and beyond the short edge we removed). This gives again $k(l_1 + l_2) - 1$ new edges, it does therefore not decrease the edge-number, but the new graph allows extension by further pointS, each giving k new edges (extending the subpath of Σ). So for those $n > n_0(T, \Gamma)$ with $h(n+1) = h(n) + (k-1)$ this cannot happen. For those n the extremal sections are generated by a polygon with T-parallel edges, and for the other n in between we can obtain extremal sections by removing points along an edge.

For those n with $h(n+1) = h(n) + (k-1)$ the edges of the extremal sets must have almost equal relative lengths: it cannot happen that we remove a path of l vertices along one edge of Z (losing $kl - 1$ graph-edges) and find another edge of Z on which there is a missing path Σ of length $l^* \geq l + 1$ (in which case we could add $l + 1$ points with $k(l+1) - 1$ new edges, contradicting the assumption on $h(n+1)$).

So the extremal sections X_n are generated by polygons with T-parallel edges of almost equal relative length (as multiple of the corresponding vector of T). So each edge has a length of $\alpha_t \sqrt{n} + \Theta(1)$. To count the number of edges in X_n, we note that outside the neighbourhoods of the $2k$ vertices of the generating polygon, we lose edges only along the edges, at a rate proportional to the length of the edge (the constant of course also depending on the direction). This gives the asymptotics claimed in the lemma

$$h(n) = kn - c_{T,\Gamma}\sqrt{n} + \Theta(1).$$

We also note that if we have a set consisting of a section of $n \geq n_0(T, \Gamma)$ vertices and any number of arbitrary further vertices, we can extend this set by any given number m of vertices in such a way that we gain at least $h(n+m) - h(n) - 1$ additional edges. To prove this, we extend at the same time the section in the given set and the optimal section and compare the numbers of edges. The points are added in several stages using $m_0 + m_1 + \cdots + m_r = m$ points, each going until either there are no further points left or we have reached a number of points $n + m_1 + \cdots + m_i$ for which h grows only by $k - 1$ (so a path along the boundary of the extremal section has been completed). We first disregard the further points outside the given section, and count only those edges going from the added points to the section. Let $m_0 = \min\{j \geq 0 \mid h(n+j+1) < h(n+j) + k\}$ and $m_{i+1}i = \min\{j \geq 1 \mid h(n + m_0 + \cdots + m_i + j + 1) < h(n + m_0 + \cdots + m_i + j) + k\}$. The extremal section of a given number of points minimizes the maximum relative length of its boundary edges, so the section of the given set always has an edge of at least the length of the edge of the extremal set. This edge may already contain some points of the section, after filling it up we can start a new edge, so we can always add the m_i points in such a way that we gain at least $km_i - 1$ new edges. For all $i > 0$ we have by definition $h(n + m_0 + \cdots + m_i) - h(n + m_0 + \cdots + m_{i-1}) = km_i - 1$. Only $h(n + m_0) - h(n) = km_0$. This proves the claim, if there are no further points. For each further point we have met in our construction, we have to add a point to the final section with as many edges going to the section as we counted for the point when we met him. We met these points as parts of edge-paths being filled up, and since the edges of the final section are at least as long as the edge-length when we met them, we can just add such paths again, which proves the claim.

Now we have to determine the structure of the extremal subsets; of these we already know that they consist of some T-parallel convex sections $Z_1, \ldots, Z_l$ and at most $O(\sqrt{n})$ further scattered points.

Let $n_i := |\Gamma \cap Z_i|$ be the number of points of S_n in Z_i. By the isoperimetric inequality we find that the boundary of Z_i has length $\Omega\left(\sqrt{n_i}\right)$, so there are at least $\Omega\left(\sqrt{n_i}\right)$ points in $Z_i \cap S_n$ which do not have the full degree in $G_T(S_n)$. Since the total number of these points in S_n is $O\left(\sqrt{n}\right)$, we have $\sum_{i=1}^{l} \sqrt{n_i} \leq \alpha\sqrt{n}$. Using $\sum_i n_i = n$ we find $\max_i n_i \geq \varepsilon n$; so there is a big connected component. Also we note that there are only $O(\sqrt{n})$ connected components, since each summand is at least one.

All those points of S that do not lie in any of the Z_i will nonetheless be near one of the Z_i, the distance will remain bounded by a constant δ. For each point of S_n that does not belong to any of the Z_i has a graph-distance at most d to a point of S_n that does not have full degree in $G_T(S_n)$. So if in the graph-metric of $G_T(S_n)$ there is a ball with a big radius ρ around a point of S_n that contains only points which do not belong to any Z_i ($\hat{n} \geq \rho$ such points), then the average degree in the interior of that ball is at most $2k - \epsilon(d)$, and the number of edges leaving the ball is at most $O\left(\sqrt{\hat{n}}\right)$. It follows $f(n-\hat{n}) \geq f(n) - (2k-\epsilon)\hat{n} - O\left(\sqrt{\hat{n}}\right)$; for sufficiently large $\hat{n}$, i.e. sufficiently large radius ρ this gives a contradiction to $f(n) \geq f(n-\hat{n}) + f(\hat{n}) \geq f(n-\hat{n}) + k\hat{n} - O\left(\sqrt{n}\right)$. So the maximum distance ρ of a point not belonging to any Z_i to the nearest point in a Z_i stays bounded: $\rho = O(1)$.

Let $Y_i \subset \mathbb{R}^2$ denote the set of points with distance at most ρ to Z_i. The union $\bigcup_{i=1}^{l} Y_i$ contains all points of S_n. Two set Y_i, Y_j can overlap, but for any set Y_i the total area overlapped by over Y_j is bounded by a constant. For if Y_i is overlapped along an edge with direction t_a by over sets Y_j with a total length λ, then we can increase S_n by all those points on parallels to that edge that lie between Z_i and Z_j. These points can be partitioned as above in arithmetic progressions with difference t_a (one-dimensional sublattices), where the number of distinct such progressions that fits between Z_i and Z_j is bounded by the distance $O(\rho) = O(1)$. In each such sequence $p, p+t_a, p+2t_a, \ldots$ the first point generates at least $k-1$ new edges, and all further points (till the end of Z_i) generate at least k new edges. But as soon as these parallels reach Z_j, each point gives at least one further edge across to Z_j, which gives $\Omega(\lambda)$ further edges. So if we added a total of μ new points, we got at least $k\mu - O(\rho) + \Omega(\lambda)$ additional edges, which gives a contradiction to $f(n+\mu) \leq f(n) + k\mu$ unless λ stays bounded by a constant.

If there is only one connected component, the claim of the Lemma follows directly. So we assume there are several components $Z_1, Z_2, \ldots$, which are ordered with decreasing size n_i. Let also $m_i := |(Y_i \setminus Z_i) \cap S_n|$. For each constant σ the claim of the proof follows for sufficiently big $n > n_0(\sigma)$, if $n_2 < \sigma$ holds, since there are only at most $O(\sqrt{n})$ components. So we can assume that n_2 gets arbitrarily big.

Suppose now that $m_2 \le \frac{1}{2} n_2$ holds. Each set Y_i contains at most $h(n_i) + km_i + O(1)$ edges: at most $h(n_i)$ within Z_i, at most km_i edges by shelling those point of $Y_i \setminus Z_i$ which are not overlapped by another Y_j, and at most $O(1)$ edges from those $O(1)$ points, in which Y_i is overlapped by other Y_j.

If we remove now the n_2+m_2 points of $S_n \cap Y_2$, and extend Z_1 by this number of points we lose at most $h(n_2)+km_2+O(1)$ edges in Y_2 and gain at least $h(n_2+m_2+n_1)-h(n_1)-1$ edges around Z_1 But using $m_2 \le \frac{1}{2} n_2$ we find

$$\begin{aligned}
\Big(h(n_2+m_2+n_1) - h(n_1) - 1\Big) - \Big(h(n_2) + km_2 + O(1)\Big) \\
&= \Big(kn_2 + km_2 - c_k\Big(\sqrt{n_2+m_2+n_1} - \sqrt{n_1}\Big) + O(1)\Big) \\
&\quad - \Big(kn_2 + km_2 - c_k\sqrt{n_2} + O(1)\Big) \\
&= -c_k \frac{n_2+m_2}{\sqrt{n_1+n_2+m_2} + \sqrt{n_1}} + c_k\sqrt{n_2} - O(1) \\
&> -c_k \frac{n_2+m_2}{2\sqrt{n_1}} + c_k\sqrt{n_2} - O(1) \\
&> \frac{c_k}{4}\sqrt{n_2} - O(1),
\end{aligned}$$

so for sufficiently big n_2 this increased the number of edges, a contradiction to the maximality of S_n.

And finally, if $m_2 > \frac{1}{2} n_2$ holds and n_2 is sufficiently big, then the set Z_2 is not 'round', but in one direction of length $\Omega(n_2)$, and orthogonal to that of width bounded by a constant. In that case a positive fraction of the points of Z_2 do not have full degree $2k$; by removing the points of Y_2 we therefore loose at most $km_2 + (k-\epsilon)n_2$ edges, and by increasing Z_1 we gain again at least

$$kn_2 + km_2 - c_k\Big(\sqrt{n_1+n_2+m_2} - \sqrt{n_1}\Big) \ge kn_2 + km_2 - O\Big(\sqrt{n_2}\Big)$$

edges ($m_2 = O(n_2)$), which generates for sufficiently big n_2 again a contradiction.

So all but $O(\sqrt{n})$ points of S_n are in the first connected component, which was shown to be a convex polygonal section of Γ bounded by edges parallel to T. Since we can remove the $m = O(\sqrt{n})$ outside points as they become available as vertices of the convex hull of the set, we have at most $h(n-m)+km$ edges; since $\sqrt{n} - \sqrt{n-m} = O(1)$, we have $f(n) = h(n) + O(1)$. But this implies that the edge-lengths of that component can differ by at most $O(1)$ from the edge-lengths of the optimal section, so the optimal subset differs from the optimal section by at most $O(\sqrt{n})$ points.
This completes the proof of Lemma 2 and the Theorem.

4. References

[1] J. Beck and J. Spencer: Unit distances, *J. Combinatorial Theory Ser. A* **37** (1984), 231–238

[2] P. Brass: Erdős distance problems in normed spaces, *Computational Geometry - Theory and Applications* **6** (1996) 195–214.

[**3**] P. Brass: Extremale Konstruktionen in der kombinatorischen Geometrie, Habilitationsschrift, Universität Greifswald 1996

[**4**] P. Brass: Häufige Abstände in endliche Punktmengen, to appear in *Jahrbuch Überblicke Mathematik*

[**5**] K.L. Clarkson, H. Edelsbrunner, L.J. Guibas, M. Sharir and E. Welzl: Combinatorial complexity bounds for arrangements of curves and spheres, *Discrete & Computational Geometry* **5** (1990), 99–160

[**6**] K.B. Chilakamarri: Unit distance graphs in Minkowski metric spaces, *Geometriae Dedicata* **37** (1991), 345–356

[**7**] K.B. Chilakamarri: The unit-distance graph problem: a brief survey and some new results, *Bulletin of the ICA* **8** (1993) 39–60

[**8**] K.B. Chilakamarri and C.R. Mahoney: Maximal and minimal forbidden unit-distance graphs in the plane, *Bulletin of the ICA* **13** (1995) 35–43

[**9**] P. Erdős: On sets of distances of n points, *Amer. Math. Monthly* **53** (1946), 248–250

[**10**] F. Harary and H. Harborth: Extremal animals, *J. Combinatorics, Information & System Sciences* **1** (1976) 1–8

[**11**] H. Harborth: Lösung zu Problem 664A, *Elemente der Mathematik* **29** (1974) 14–15

[**12**] S. Józsa and E. Szemerédi: The number of unit distances in the plane, 939–950 in *Infinite and Finite Sets* (A. Hajnal e.a., Eds.) North Holland, Amsterdam, 1975

[**13**] J. Pach and P.K. Agarwal: *Combinatorial Geometry*, Wiley, New York 1995

[**14**] C. Purdy and G. Purdy: Minimal forbidden distance one graphs, *Congressus Numerantium* **66** (1988) 165–172

[**15**] C. Schade: *Exakte Maximalzahlen gleicher Abstände*, Diplomarbeit bei H. Harborth, Universität Braunschweig 1993

[**16**] J. Spencer, E. Szemerédi and W. Trotter: Unit distances in the euclidean plane, 293–304 in *Graph Theory and Combinatorics* (B. Bollobás, Ed.) Academic Press, London, 1984

[**17**] L.A. Székely: Crossing numbers and hard Erdős problems in discrete geometry, Preprint 1996